Investigating Biological Approaches for Monitoring and Manipulating Microbial Communities in Anaerobic Digestion

Nathan Innard

PhD

University of York

Biology

December 2025

Abstract

Anaerobic digestion (AD) is green technology used to treat anthropogenic biological wastes while simultaneously generating renewable energy. Our understanding of the composition of the mixed microbial communities that underpin AD has increased substantially in recent years thanks largely to metagenomic sequencing efforts, yet our ability to monitor and manipulate them in a controlled and predictable manner has not kept pace. The objective of this project was to investigate biological approaches at the whole organism level through which this might be achieved. Methods aimed at improving AD whole cell biosensors through integration of volatile reporter genes were investigated. Genetic techniques were used to exogenously express a recently discovered anaerobic ethylene synthesis pathway, with results demonstrating preliminary evidence of successful synthesis of this valuable chemical precursor in an engineered bacterial system. Proof of principle work was used to demonstrate the potential of targeted microbial isolation techniques and directed evolution for AD chassis organism improvement, resulting in the successful generation of bacterial strains with increased fitness in anaerobic conditions, accompanied by sequencing data revealing possible causative genetic changes. Culturomics and a custom isolate identification strategy were also successfully applied to AD microorganisms. It is hoped that the results and ideas presented in this thesis will facilitate further development of targeted AD monitoring and manipulation approaches in the future.

Table of Contents

Abstract 2

Table of Contents 3

List of Figures 13

List of Tables 16

Acknowledgements 17

Declaration 18

Chapter 1: General Introduction 19

1.1 The need for sustainable technologies 19

1.2 Anaerobic digestion 19

1.3 The utility of AD 19

1.4 Global implementation of AD 20

1.5 AD for sewage treatment 20

1.6 Pretreatment 20

1.7 The AD process 21

1.8 The AD community 22

1.9 AD community stability 22

1.10 Identifying microbes in mixed microbial communities 22

1.11 DNA sequencing for microbial identification 23

1.12 Classification and species boundaries considerations 24

1.13 AD community monitoring 25

1.14 Existing AD monitoring techniques 26

1.14.1 Physicochemical monitoring 26

1.14.2 Titration based monitoring 26

1.14.3 Spectroscopic monitoring 27

1.14.4 Chromatographic monitoring 27

1.14.5 Molecular biology techniques 27
1.15 AD monitoring deployment 28
1.16 Engineered whole cell biosensors for AD community monitoring 29
1.17 Existing whole cell biosensors 29
1.18 Reporter genes 30
1.19 Whole cell volatile biosensors 30
1.20 Existing whole cell volatile reporter systems 31
1.21 The importance of the microbial chassis 32
1.22 Existing chassis organisms 32
1.23 AD community manipulation 33
1.23.1 Bioaugmentation 34
1.23.2 Synthetic mixed communities 34
1.23.3 Use of phage 35
1.23.4 Microbial genetic modification 36
1.24 Aims 37
Chapter 2: Methods 39
2.1 Molecular work 39
2.1.1 Construct design 39
2.1.2 Codon optimization 39
2.1.3 Gene synthesis 39
2.1.4 Plasmid material 39
2.1.5 Bacterial strains 39
2.1.6 Protein alignments 40
2.1.7 Re-dissolving plasmid material shipped on filter paper 40
2.1.8 Plasmid sequencing 40
2.1.9 Media preparation 40
2.1.9.1 LB broth 40

2.1.9.2 Nutrient broth 40
2.1.9.3 Premixed broth 40
2.1.9.4 IPTG 41
2.1.9.5 X-gal 41
2.1.9.6 Antibiotics 41
2.1.9.7 LB agar 41
2.1.9.8 Schaedler Agar 41
2.1.9.9 Premixed solid media 42
2.1.9.10 Agar plates 42
2.1.9.11 Premixed diluents 42
2.1.9.12 Resazurin Solution 42
2.1.9.13 Anaerobic LB broth supplemented with resazurin 42
2.1.9.14 Anaerobic chloramphenicol solution 43
2.1.10 *E. coli* culturing 43
2.1.11 Glycerol stocks 43
2.1.12 Generating chemically competent cells (*E. coli*) 44
2.1.13 Plasmid DNA isolation 44
2.1.14 DNA quantification 44
2.1.15 *E. coli* transformation 45
2.1.16 Diagnostic plasmid digestion 45
2.1.17 Golden Gate assembly reactions 45
2.1.18 Agarose gel electrophoresis 45
2.1.19 Genomic DNA extractions 46
2.1.19.1 From sludge 46
2.1.19.2 From microbial cell cultures 46
2.1.19.3 From platewashes 46
2.1.19.4 For kit comparison for 8 organism test pool experiment 46

2.1.19.5 For 96 organism test array pooling 47
2.1.20 Sludge sampling 48
2.1.21 Sludge plating 48
2.1.22 Colony counting and colony forming unit determination 48
2.1.23 Experimental evolution 49
2.1.23.1 Mixed strain plate comparisons 49
2.1.23.2 Resazurin impact experiment 49
2.1.23.3 Evolution passage 50
2.1.23.4 Comparative growth experiment 50
2.1.24 Selected organism isolation on plates 51
2.1.24.1 *E. coli* 51
2.1.24.2 *B. subtilis* 51
2.1.24.3 *C. beijerinckii* 52
2.1.24.4 Colony selection and liquid culture outgrowth 52
2.1.25 Separately autoclaved media plating 52
2.1.25.1 R medium 52
2.1.25.2 T and S medium 53
2.1.26 Colony picker work 53
2.1.26.1 Colony picking and inoculation into liquid culture 53
2.1.26.2 *E. coli* isolation and outgrowth test 54
2.1.26.3 *E. coli* glycerol storage and revival tests 54
2.1.27 Ethylene sampling and analysis (photoionisation detector) 55
2.1.28 Sludge viability experiment 55
2.1.28.1 Sludge starving 55
2.1.28.2 BMP test 56
2.1.28.3 Sludge storage 57
2.1.29 Methane quantification 57

2.2 GC-MS sampling and analysis 57
2.2.1 Originally used GC-MS 57
2.2.1.1 Acetone plasmid growth experiment sampling 57
2.2.1.2 Acetone/isoprene spiking experiment sampling 58
2.2.1.3 Volatile analysis 58
2.2.2 TF GC-MS 59
2.2.3 GC used for ethylene 59
2.2.3.1 Calibrations 59
2.2.3.2 Initial ethylene production test 60
2.2.3.3 Varied antibiotic concentration ethylene test 61
2.2.3.4 DH5α ethylene test 61
2.2.3.5 Repeat test of putative ethylene producing cultures 61
2.2.3.6 Running headspace samples on GC 61
2.2.3.7 Peak integration 62
2.3 Sequencing and bioinformatics 62
2.3.1 16S metagenomic sequencing 62
2.3.2 Identification of desirable organisms for isolation from sludge 62
2.3.2.1 BONCAT metagenomic data 62
2.3.2.2 Long term digester metagenomic data 63
2.3.2.3 Assembly of metagenomic data into MAGS and identity assignment 63
2.3.2.4 Generating list of “wanted” organisms 63
2.3.2.5 Generating and querying custom BLAST database: genomes 64
2.3.2.6 Generating and querying custom BLAST database: 16S 64
2.3.3 Pre- and post-THP platewash 16S sequencing and analysis 65
2.3.4 Eight organism test pool sequencing, assembly and analysis 66
2.3.5 96 well plate sequencing and analysis 66
2.3.6 Selective plating isolate whole genome sequencing and identification 67

2.3.7 Experimental evolution sequencing and SNP calling 67
2.4 Miscellaneous 68
2.4.1 Figures 68
2.4.2 Statistics 68
Chapter 3: Whole Cell Volatile Bioreporters 69
3.1 Introduction 69
3.1.1 Context of work 69
3.1.2 Repurposing an existing system 69
3.1.3 Identifying additional candidate volatiles 69
3.1.4 Anaerobic ethylene synthesis 70
3.1.5 Aims 71
3.2 Results 72
3.2.1 Select additional candidate volatiles which can be biosynthesised anaerobically 72
3.2.1.1 Rational exclusion of volatiles already detectable on initially available GC-MS. 72
3.2.1.2 Additional volatiles not already detectable on available GC-MS 73
3.2.2 Obtain plasmid constructs suitable for producing selected volatiles anaerobically 74
3.2.2.1 Ethylene and 2(methylthio)ethanol 74
3.2.2.1.1 Pathway engineering rationale 74
3.2.2.1.2 Obtaining genes 74
3.2.2.1.3 Assembly 75
3.2.2.2 Acetone 79
3.2.2.3 Other volatiles 81
3.2.3 Validate biosynthesis of chosen volatiles in *E. coli* 81
3.2.3.1 Acetone 81
3.2.3.2 Isoprene 86

3.2.3.3 Trying a different GC-MS 87
3.2.3.4 Testing the remaining four volatiles on the TF GC-MS 87
3.2.3.5 Refocusing aims to anaerobic ethylene synthesis 90
3.2.3.5.1 Ethylene detection by infrared sensor 91
3.2.3.5.2 Ethylene detection by GC 91
3.2.3.6 Testing ethylene producing strains 92
3.2.3.6.1 Generating strains 92
3.2.3.6.2 Ethylene production tests 94
3.2.3.6.3 Initial ethylene production test 94
3.2.3.6.4 Varied antibiotic concentration ethylene test. 100
3.2.3.6.5 DH5α ethylene test 103
3.2.3.6.6 Repeat testing of putative ethylene producing cultures 105
3.3 Discussion 107
3.3.1 Aims recap 107
3.3.2 Challenges encountered 107
3.3.3 Chapter achievements 107
3.3.4 Was the ethylene putatively detected genuine? 108
3.3.5 Ethylene production from the control plasmid 109
3.3.6 Why could the putative ethylene synthesis result not be reproduced? 109
3.3.7 Possible future work for this arm of project 110
3.3.8 Next steps 112
Chapter 4: Chassis Organisms 113
4.1 Introduction 113
4.1.1 Context of work 113
4.1.2 Desirable features of AD chassis organisms 113
4.1.3 Potential strategies to increase robustness of chassis organisms in AD 113
4.1.4 Directed evolution 114
4.1.5 Isolating endogenous organisms from an AD community 115

4.1.6 Aims 116
4.2 Results 117
4.2.1 Chassis organisms: directed evolution 117
4.2.1.1 Obtaining suitable organisms 117
4.2.1.2 Testing impact of resazurin/validation work 120
4.2.1.3 Evolving strains 122
4.2.1.4 Fitness assay 123
4.2.1.5 Strain sequencing 125
4.2.2 Obtaining chassis organisms: plating 130
4.2.2.1 Selective plating from sludge 130
4.2.2.1.1 *In silico* identification of candidate organisms 130
4.2.2.1.2 Attempted isolation of chosen organisms from sludge 130
4.2.2.2 Untargeted isolation of organisms from sludge 137
4.2.2.2.1 Simple sludge plating conditions 137
4.2.2.2.2 Broadening sludge plating conditions 139
4.2.2.2.3 Recovery of unidentified isolated organisms 142
4.2.3 Identifying isolated organisms 145
4.2.3.1 Method development: single 8 organism test pool 146
4.2.3.2 96 Sample pooling 155
4.3 Discussion 162
4.3.1 Aims recap 162
4.3.2 Chapter achievements 162
4.3.3 Directed evolution 162
4.3.4 Selective plating of sludge 167
4.3.5 Non-selective plating of sludge 169
4.3.6 Sequence based microorganism identification 170
4.3.7 Next steps 171
Chapter 5: Culturomics 172

5.1 Introduction 172
5.1.1 Context of work 172
5.1.2 Possible advantages of a simplified AD community 172
5.1.3 An AD SynCom as a simplified community 172
5.1.4 Culturomics 173
5.1.5 Working from consistent communities 173
5.1.6 Storing mixed microbial communities 174
5.1.7 Plating of anaerobes in aerobic conditions 174
5.1.8 Aims 175
5.2 Results 176
5.2.1 Investigating system feed 176
5.2.1.1 Plating Esholt pre- and post- THP feed 176
5.2.1.2 16S DNA analysis of pre- and post- THP feed 178
5.2.1.3 DNA analysis of pre- and post- THP plate washes 183
5.2.2 Can a viable AD community be stored? 184
5.2.2.1 Developing a reliable methane sensor 185
5.2.2.2 Stored community viability test 186
5.2.2.3 Community composition changes during storage 187
5.2.3 Colony picker pipeline development 189
5.2.3.1 Isolation and liquid culture outgrowth 190
5.2.3.2 Storage and revival 192
5.2.4 Anaerobic plates 192
5.2.4.1 Separate autoclaving with antioxidants 192
5.2.4.2 Separate autoclaving only 194
5.2.5 High throughput colony isolation from stored sludge 200
5.3 Discussion 203
5.3.1 Aims Recap 203

5.3.2 Chapter achievements 203
5.3.3 System feed investigations 203
5.3.4 Sludge storage community changes 205
5.3.5 Anaerobic plates 206
5.3.6 Colony picker protocol 207
5.3.7 Further work: generating and testing an AD library 207
Chapter 6: General Discussion 210
6.1 Aims recap 210
6.2 Achievements and challenges 210
6.3 The potential of combined approaches 210
6.4 Immediate further work 211
6.5 Longer term further work 212
Appendix: 216
Reference List 226

List of Figures

Figure 1.1: Core components of a whole cell microbial biosensor. 29

Figure 1.2: Potential advantages of volatile over visual reporters in AD. 31

Figure 3.1: Ethylene synthesis pathways. 71

Figure 3.2: O-acetylhomoserine sulfhydrylase (O-as) alignments. 75

Figure 3.3: 2(methylthio)ethanol synthesis plasmids. 76

Figure 3.4: Ethylene synthesis plasmids (partial pathway). 77

Figure 3.5: Ethylene synthesis plasmids (full pathway). 78

Figure 3.6: Acetone synthesis plasmids. 80

Figure 3.7: Overview of acetone plasmid validation experiment. 82

Figure 3.8: Growth dynamics of *E. coli* carrying pETM6_Ptrc. 83

Figure 3.9: Relative acetone readings for *E. coli*. 84

Figure 3.10: Overview of acetone spiking experiment. 85

Figure 3.11: Relative acetone readings following spiking. 86

Figure 3.12: Relative isoprene readings following spiking. 87

Figure 3.13: Representative traces generated on TF GC-MS. 89

Figure 3.14: Ethylene calibrations on GC. 92

Figure 3.15: Overview of initial ethylene production test. 95

Figure 3.16: Growth dynamics and headspace ethylene readings - initial ethylene test. 97

Figure 3.17: Representative ethylene traces generated on GC. 99

Figure 3.18: Varied antibiotic concentration ethylene test. 102

Figure 3.19: DH5α ethylene test. 104

Figure 3.20: Repeat of initial culture ethylene test. 106

Figure 4.1: Possibilities for obtaining or generating more robust AD chassis organisms. 114

Figure 4.2: Appearance of mixed fluorescent and non-fluorescent *E. coli* strains. 118

Figure 4.3: Pairwise *E. coli* strain mixing colony counts. 119

Figure 4.4: Effect of resazurin on *E. coli* culture turbidity readings. 121

Figure 4.5: Directed evolution approach. 122

Figure 4.6: Directed evolution passaging. 123

Figure 4.7: Comparative growth curves. 125

Figure 4.8: Selective plating overview. 131

Figure 4.9: Colony counts of selective vs non-selective plates. 132

Figure 4.10: Selective plate controls. 133

Figure 4.11: Taxonomic classification of short read selective plate DNA samples. 135

Figure 4.12: Plated sludge colony morphologies and counts over time. 138

Figure 4.13: Aerobic plated sludge colony morphologies and counts between media. 140

Figure 4.14: Anaerobic plated sludge colony morphologies and counts between media. 142

Figure 4.15: Microbial recovery success rates following cryostorage. 144

Figure 4.16: Proposed organism identification strategy. 145

Figure 4.17: Morphologies of organisms A - H. 146

Figure 4.18: Organism A-H identification based on 16S sequencing. 147

Figure 4.19: Taxonomic classification of unassembled “pure” DNA samples A and D. .. 149

Figure 4.20: Taxonomic classification of unassembled pooled DNA samples. 151

Figure 4.21: Classification of binned individual samples. 153

Figure 4.22: Classification of binned pooled samples. 154

Figure 4.23: Array layout for 96 sample pooling experiment. 157

Figure 4.24: Identities for 96 sample pooling experiment. 161

Figure 5.1: Colony counts of pre- and post-THP plated feed. 177

Figure 5.2: Colony morphologies of pre- and post-THP plated feed. 178

Figure 5.3: Number of ASVs in different conditions. 179

Figure 5.4: Taxonomic classification of abundant groups between conditions. 182

Figure 5.5: Sludge storage and viability testing process. 185

Figure 5.6: Gas production following sludge storage. 187

Figure 5.7: Observed features by sample. 188

Figure 5.8: Taxonomic classification of abundant genera. 189

Figure 5.9: *E. coli* test array plate layout. 190

Figure 5.10: *E. coli* liquid culture outgrowth. 191

Figure 5.11: Colony recovery of sludge on supplemented and control plates. 193

Figure 5.12: T and S medium comparison (aerobic incubation). 195

Figure 5.13: T and S medium comparison (anaerobic incubation). 196

Figure 5.14: T and S medium plate wash 16S analysis. 199

Figure 5.15: Plated stored sludge colony morphologies and counts between media.... 201

Figure 5.16: Proposed approach for generation AD SynComs. 208

Figure 6.1: An example of combined approaches for AD community manipulation. 211

Figure A1: Ethylene and 2(methylthio)ethanol producing plasmid assembly scheme. . 220

Figure A2: Acetone calibrations. 222

Figure A3: Initial ethylene test media controls. 223

Figure A4: Ethylene headspace repeat readings. 223

Figure A5: CFU to OD_{600} calibrations for *E. coli* RFP strain. 224

Figure A6: Reproducible methane content measurements with an IR sensor. 224

Figure A7: Irregularly shaped colonies on T and S plates. 225

List of Tables

Table 3.1: Suitability of existing GC-MS volatiles for anaerobic gas reporting. 73

Table 3.2: Proposed *E. coli* strain/gene combinations for ethylene synthesis. 74

Table 3.3: Detectability of volatiles of interest at different dilutions on TF GC-MS. 90

Table 3.4: Plasmid status in *E. coli* cultures following ethylene synthesis tests. 105

Table 4.1: Mutant allele frequencies in experimentally evolved *E. coli* populations. 126

Table 4.2: Identification of selective plating DNA samples. 136

Table 4.3: Summary of contents of unassembled individual DNA samples A-H..... 152

Table 4.4: Comparison of organism identification between different methods. 159

Table 5.1: Alpha diversity metrics for different conditions. 180

Table 5.2: Alpha diversity metrics for plate was 16S data. 197

Table 5.3: High throughput liquid culture recovery of picked colonies..... 202

Table A1: Species identified for investigation..... 216

Table A2: List of species identified as “desirable” organisms..... 219

Acknowledgements

I would like to thank my supervisors James Chong, Matthew Reilly, and Daniela Barillà for their guidance and support throughout this project. I would also like to thank Sarah Forrester who has provided support on the level of supervision throughout my project. Thanks to all members of the Chong lab past and present for their companionship and good humour and particularly to Annabel Cansdale for additionally helping me with bioinformatics.

I am grateful to my family for all their support throughout my PhD: my brothers Alex, Aaron and Joel for being so understanding and my mum Poppy and auntie Angela who were always happy to help, particularly in difficult times.

Special thanks go to Emily Churchill for putting up with me throughout this process.

Declaration

I declare that this thesis is a presentation of original work and I, Nathan Innard, am the sole author. This work has not previously been presented for a degree or other qualification at this University or elsewhere. All sources are acknowledged as references.

All of the work in the thesis is my own with the following exceptions:

- Work on the TF GC-MS instrument was carried out with the assistance of Swen Langer. Optimisation to separate acetone from mixed samples was performed by Swen Langer.
- Production of the 3D printed photoionisation detector housing was carried out by Mark Bentley.
- Generation of the BONCAT dataset was carried out by Yumechris Amekan. Sampling for the standard metagenomic dataset from Naburn was performed by James Chong and Anna Alessi. Sequence assembly and bioinformatic analysis for the wanted organism identification were carried out by Annabel Cansdale.
- 16S sequencing of pre- and post-THP feed plate washes including the additional DNA amplification step was performed by the Technology Facility at the University of York. Analysis of the sequencing data generated was performed by Annabel Cansdale.
- Assembly and binning of short reads for identification in the 8 organism test pool, as well as taxonomic assignment was done by Sarah Forrester. Taxonomic assignment for the short read data from the selective plating was done by Sarah Forrester.
- Long read Nanopore sequencing of the pooled 96 sample array was performed by the Technology Facility at the University of York. Analysis of the sequencing data including generating and using the custom bioinformatics pipeline for assigning assemblies to row/column intersections and taxonomic assignment of assemblies at different levels of confidence was carried out by Annabel Cansdale.
- Some of the material in the introduction sections has been adapted from a review I co-authored (Innard and Chong 2022).

Chapter 1: General Introduction

1.1 The need for sustainable technologies

The human population is expected to rise to 9.7 billion by 2050 (Garcia *et al.* 2020). This increase will be accompanied by growing energy demands (Hickel *et al.* 2021) and the need to process ever greater amounts of waste (Hoornweg *et al.* 2013). Historically, our energy needs have been met primarily by the burning of fossil fuels, which has had a variety of negative global impacts including climate change (Atelge *et al.* 2020). Waste management has relied heavily on unsustainable methods like landfilling (Allen 2001). There is a growing appreciation that this linear system based on resource use and disposal cannot continue, and a transition to a circular economy grounded in reuse and energy recovery is required (Clark 2019). The adoption of sustainable technologies such as anaerobic digestion (AD) will be key to achieving this transition (Ibarra-Esparza *et al.* 2023; Alengebawy *et al.* 2024).

1.2 Anaerobic digestion

AD is a natural biological process in which organic materials are broken down by microorganisms in oxygen free environments (Chen *et al.* 2008). The main outputs of this process are biogas (a mixture of methane and carbon dioxide that can be used as fuel (Li *et al.* 2019)) and digestate (which can be used as fertiliser (Möller and Müller 2012)). The process is used widely in the treatment of anthropogenic wastes from a variety of sources including residues and manures from agriculture (Alengebawy *et al.* 2024), food waste (Morales-Polo *et al.* 2018) and sewage sludge (Lim *et al.* 2020).

1.3 The utility of AD

AD is particularly appealing among sustainable technologies (Subbarao *et al.* 2023) for a number of reasons. Firstly, it generates energy while simultaneously removing waste products that would need to be dealt with anyway (DelaVega-Quintero *et al.* 2025). Another benefit is the option to upgrade biogas to biomethane, a fuel compatible with existing energy infrastructure (Cansdale *et al.* 2023). Unlike intermittent renewable energy sources like wind and solar power, energy generation through AD is not weather dependent (Shinde *et al.* 2023). AD also captures greenhouse gases which would otherwise be released into the atmosphere (Zaks *et al.* 2011) during decomposition.

1.4 Global implementation of AD

AD is a well established technology and is deployed at an industrial scale for waste treatment, particularly in Europe (Atelge *et al.* 2020; Ibarra-Esparza *et al.* 2023) where Germany is world leading in terms of biogas production (M. Yang *et al.* 2024). The considerable investment required for full-scale plant installation and operation can be more of a barrier in developing countries, where small (e.g. household) scale systems are common (Ibarra-Esparza *et al.* 2023). The UK has a mature AD sector with around 700 AD plants, of which approximately one quarter are used for sewage treatment (Gustafsson *et al.* 2024).

1.5 AD for sewage treatment

AD is used widely as part of the sewage treatment process in the UK (Harley-Nyang *et al.* 2022), with around 90% of material being treated using this method (Gustafsson *et al.* 2024). Following a preliminary step to remove large objects, grit and sand (Canziani and Spinosa 2019), sewage entering a wastewater treatment works (WWTW) is passed to a settlement tank which splits the wastewater into two streams: primary sludge which is mainly organic matter that settles at the bottom of the vessel through gravity and a liquid fraction that is sent for aerobic biological treatment (Appels *et al.* 2008; Collivignarelli *et al.* 2019). This aerobic treatment produces an excess of microbial biomass known as secondary or waste-activated sludge (Canziani and Spinosa 2019; Collivignarelli *et al.* 2019). Primary and secondary sludge are then typically combined and serve as the starting material (or feed) for the AD process (Appels *et al.* 2008; Zahan *et al.* 2016). Anaerobic digesters are essentially large airtight tanks in which this material is treated, generally with continuous stirring under mesophilic temperature conditions (around 30 - 38 °C) (Appels *et al.* 2008). The sewage treatment process as a whole is very energy expensive, with electricity accounting for an estimated 25 – 50% of the running costs of a typical facility, making the energy AD generates important for offsetting these costs (Gandiglio *et al.* 2017).

1.6 Pretreatment

Influent feed may be pre-treated in a variety of different ways including through thermal, mechanical, chemical and biological methods (Carrère *et al.* 2010). The aim of these approaches is generally to increase the digestibility of the feed which in turn can increase

AD efficiency and biogas generation (Khanh Nguyen *et al.* 2021). For example, thermal hydrolysis pretreatment (THP) is an established technique through which heat and pressure are used to improve the biodegradability of sludge (Barber 2016), which is largely achieved by the destruction of cell membranes and the resultant release of organic contents (Ariunbaatar *et al.* 2014). Commercial THP processes such as Cambi and BioThelys™ are available and have been applied at process scale in working WWTW (Carrère *et al.* 2010; Ariunbaatar *et al.* 2014), including in the UK (Barber 2016). While the primary aim of thermal pretreatment is to improve sludge biodegradability, the associated cell destruction also comes with the added advantage that it destroys pathogens present in the sludge (Pilli *et al.* 2015; Barber 2016).

1.7 The AD process

The AD process itself can broadly be split into four separate consecutive stages (Venkiteshwaran *et al.* 2015). Hydrolysis, which is the first and rate limiting step of AD in the case of sewage sludge (Wang *et al.* 2022), involves the breakdown of large insoluble components like fats, proteins and polysaccharides in the feed into smaller soluble substrates (Guo *et al.* 2021). The next step is acidogenesis, where these simple units are processed to form a mixture of volatile fatty acids (VFAs), alcohols and ketones which are in turn processed into acetate, carbon dioxide and hydrogen in the acetogenesis step (Uddin and Wright 2023). The final step of the process is methanogenesis, where the products of acetogenesis are used to form methane and carbon dioxide (Goswami *et al.* 2016).

Methanogenesis can only be performed by a special group of archaea known as methanogens. The process itself has three different subtypes, hydrogenotrophic, acetoclastic and methylotrophic, which are carried out by different types of methanogens (Venkiteshwaran *et al.* 2015). Notably, the genus *Methanosarcina* is capable of all three (Dyksma *et al.* 2020). Crucially the AD process is carried out by an interconnected mixed microbial community and is reliant on different portions of the community performing different metabolic roles (Meegoda *et al.* 2018; Cao *et al.* 2025).

1.8 The AD community

Microbial communities responsible for AD are inherently complicated (Dueholm *et al.* 2024), containing potentially hundreds of even thousands of different taxa (Hassa *et al.* 2021; Mercado *et al.* 2022). These communities also differ in time and space (i.e. between different digesters, and within digesters over time (Lee *et al.* 2012; Calusinska *et al.* 2018)), further complicating the situation. Yet while composition in terms of dominant taxa can vary dramatically, composition in terms of functional guilds (i.e. groups of organisms that perform the same function within an ecosystem (Reynolds *et al.* 2023)) tends to remain fairly constant in AD (Cao *et al.* 2025; Keating *et al.* 2025). There is generally a considerable degree of redundancy within these guilds in AD systems (Vanwonterghem *et al.* 2014a; De Vrieze *et al.* 2017).

1.9 AD community stability

AD communities are sensitive to unfavorable abiotic conditions (Elsayed *et al.* 2024). Disturbances to the balance of these communities in response to challenging conditions can impede the AD process or cause it to fail entirely, necessitating efforts to reestablish the consortium which can be costly (Lim *et al.* 2020; Innard and Chong 2022), particularly at process scale. A well documented example is digester acidification, where factors such as overfeeding (Franke-Whittle *et al.* 2014) result in excessive growth of acidogenic bacteria triggering an overproduction of VFAs coupled with drops in pH which can inhibit or completely arrest methanogenesis (Z. Wang *et al.* 2023). Digester foaming, which can be caused by overgrowth of filamentous microorganisms can also result in considerable process disturbance through overflow and blockage of pipework (Jiang *et al.* 2021a). Redundancy within functional guilds provides some protection against fluctuations in abiotic conditions to the AD community on the basis of contingency (i.e. the more members a guild has, the higher the likelihood that one will be resistant to the challenging condition) (Yachi and Loreau 1999; Li *et al.* 2024; Cross *et al.* 2025). However, this vast complexity in the microbial communities that underpin AD makes predicting outcomes and disentangling cause and effect challenging (Innard and Chong 2022).

1.10 Identifying microbes in mixed microbial communities

An integral first step in understanding the AD community is knowing what organisms are there, and there are a variety of methods available for achieving this. Identification based

on morphology can be limited by the fact that different organisms can look similar on agar media, and the same organism can look different dependant on conditions (Sousa *et al.* 2013). Biochemical tests which give a profile based on different metabolic capacities of microorganisms are low throughput and labour intensive (Roy *et al.* 2023). Matrix assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF MS) is a more modern method which involves identification based on ionisation of proteins from an unidentified bacterium to generate a peptide mass fingerprint spectrum which can be cross referenced with other known spectra in database to assign identity (Singhal *et al.* 2015). MALDI-TOF for bacterial identification is both rapid and high throughput (Ashfaq *et al.* 2022). However, the need for access to both spectral databases which are often proprietary (Seuylemezian *et al.* 2018), and specialist equipment means this approach can be very expensive to establish (Singhal *et al.* 2015). The main weakness of all of these methods is that they generally require growth of bacteria as pure cultures of colonies (Braga *et al.* 2013; Arbefeville *et al.* 2024) which is limiting as the majority of microorganisms known to be present in AD have not yet been cultured (Nobu *et al.* 2020).

1.11 DNA sequencing for microbial identification

DNA sequencing as a method of identification overcomes these issues because it can be performed on mixed bacterial samples and unculturable organisms (Schloss and Handelsman 2005). Taxonomic identification can be achieved by sequencing the ubiquitous ca. 1500 base pair (bp) 16S ribosomal RNA gene (Tringe and Hugenholtz 2008; Bukin *et al.* 2019). This consists of nine hypervariable regions which aid classification, separated by conserved regions which enable PCR amplification prior to sequencing using conserved primers (Matsuo *et al.* 2021). Sequencing of the full gene is possible but partial (amplicon) sequencing of hypervariable regions V3 – V4 is often used for classification of environmental samples (Santos *et al.* 2020) including in AD (De Vrieze *et al.* 2018; Hassa *et al.* 2021). When used on mixed samples 16S sequencing can give an idea of community composition (Curry *et al.* 2022). Historically molecular operational taxonomic unit (OTU) based methods were commonly used to first cluster reads at a defined similarity threshold (typically 97%) prior to taxonomic assignment (Callahan *et al.* 2017). This has now largely been superseded by amplicon sequence variant techniques (ASVs) which do not involve arbitrary clustering and instead resolve sequences down to differences of a single nucleotide (Callahan *et al.* 2017; Chiarello *et al.* 2022). This type of sequencing comes with

the advantage that it is relatively low cost but a major limitation is that it offers lower taxonomic resolution than shotgun metagenomics (Peterson *et al.* 2021).

Whole genome metagenomic shotgun sequencing involves untargeted sequencing of DNA fragments from all microbes in a sample (Quince *et al.* 2017). Bioinformatic tools are used to assemble reads based on overlaps into progressively longer stretches of sequence called contigs and scaffolds (Mallawaarachchi *et al.* 2024). Further computational tools can then be used to collect contigs with shared features into a series of "bins" or metagenome assembled genomes (MAGS) (Mallawaarachchi *et al.* 2024) which represent a reconstruction of the microbial genomes present within the sample based on the sequencing data generated (Setubal 2021). Advantages of whole genome metagenomic sequencing over 16S include greater taxonomic resolution down to the species (Ranjan *et al.* 2016) or even strain (Snipen *et al.* 2021) level. The drawback is that it is more expensive than 16S sequencing (Hilton *et al.* 2016), requires more complex data analysis (Teeling and Glöckner 2012), access to high performance computing resources and use of software generally only compatible with Linux systems (Liu *et al.* 2021).

1.12 Classification and species boundaries considerations

Even once sequence information has been obtained it is not always straightforward to determine taxonomic composition of a mixed sample. Firstly, the principle of what defines a different microbial species is not straightforward (Achtman and Wagner 2008). The degree of similarity between the genomes of two microbes can be estimated by metrics like average nucleotide identity (ANI), but while it has been suggested that a similarity value of over 95% provides a good estimate of species level similarity (Jain *et al.* 2018), there is no universal agreement on a set cut off value (Murray *et al.* 2021; Setubal 2021). Similar issues exist in 16S sequencing, for example it is possible for two different species to have the same ASV (Strube 2021). Even strains of the same species are likely to vary at the DNA level (Martínez-Carranza *et al.* 2018), and horizontal gene transfer (HGT) in microbes further blurs these boundaries (Eisen 2000). Secondly, different sequencing methods and bioinformatic tools for assembly and binning can introduce biases or errors giving different outcomes depending on the tools used (Allali *et al.* 2017; Sczyrba *et al.* 2017; Prodan *et al.* 2020). Finally the genome reference database itself is important as classification between them can vary (Smith *et al.* 2022) and is further complicated in the

case of MAGS which may be incompletely annotated in reference databases (Lindner and Renard 2015). As a result of this, classification based on sequencing is imperfect and should be seen as a best estimate based on the information available.

1.13 AD community monitoring

The ability to monitor the health of the complex mixed microbial communities responsible for AD could be used as an early warning system for future destabilisation events, potentially avoiding costly outcomes (Wu *et al.* 2019; Lim *et al.* 2020; Innard and Chong 2022). In extreme cases, digester failure can result in the process needing to be restarted (Lim *et al.* 2020). This may result in higher operational costs (Basak *et al.* 2021), financial losses from reduced biogas production (Moeller and Zehnsdorf 2016), and even possible environmental damage (Basak *et al.* 2021). Restarting the AD process cannot be achieved instantaneously, and it may take months to reestablish the mixed microbial community that underpins the process (Lim *et al.* 2020; Santinello *et al.* 2024) so these costs may need to be borne for a considerable period of time. Direct costed examples of digester disruptions are difficult to find in the literature, but in one documented example the cost of damage caused to a digester roof following an extreme foaming event at a facility in Germany was estimated at €500,000 (Moeller and Görsch 2015). Similarly, a 10 week foaming event at a WWTW in 1996 in Sweden caused a 40% reduction in biogas generation which cost an estimated $150,000 in lost electricity generation and increased staff and consumable costs (Westlund *et al.* 1998).

Such events are so undesirable that AD plants are generally deliberately underfed to reduce the risk of process failure (Lim *et al.* 2020), and suboptimal AD operations are likely to be the norm rather than the exception (Nielsen and Angelidaki 2008; Kougias *et al.* 2014; Lima *et al.* 2023; Harirchi *et al.* 2026). This means effective AD process monitoring would not only be useful for avoiding process failure but could also enable plants to be run closer to their actual capacity, which would improve the efficiency of the process (Madsen *et al.* 2011). An ideal AD monitoring system would be cheap, easy to operate, and capable of monitoring parameters predictive of community health/function in real-time (Cruz *et al.* 2021). Such a system would ideally also be in-line (i.e. integrated into the digester or waste stream itself therefore not requiring sample removal (Madsen *et al.* 2011; Gorla *et al.* 2026)).

1.14 Existing AD monitoring techniques

A variety of intrinsic AD process parameters such as temperature and pressure, as well as those which report on products and intermediates of the system such as gas production rate or VFA composition can be monitored (L. Li *et al.* 2018). The concept is that these parameters can give an indication of how stably the process is running and the health of the underlying microbial community (Ibarra-Esparza *et al.* 2025). For example, since VFAs are intermediate products of AD and are processed relatively quickly in a balanced system, a build up can indicate digester overloading and/or methanogen inhibition (Madsen *et al.* 2011) which could precede digester acidification (see section 1.9). Various techniques are available for monitoring these parameters, and many parameters can be measured with multiple different approaches (Spanjers and van Lier 2006; Cruz *et al.* 2021).

1.14.1 Physicochemical monitoring

Numerous different methods for measuring basic operational parameters such as temperature, flow rate (liquid and gas), pH and pressure are available (Spanjers and van Lier 2006; Madsen *et al.* 2011). For example, gas flow rate can be measured using various methods including tipping bucket systems based on liquid displacement (Walker *et al.* 2009), mass flow meters (Yang *et al.* 2022) and rotameters (Spanjers and van Lier 2006). The advantage of basic physicochemical monitoring approaches is that they are generally well established with sensors commercially available (Ibarra-Esparza *et al.* 2025), and are straightforward to implement in-line (Spanjers and van Lier 2006). A key disadvantage is that they provide limited information that is of predictive value for AD community instability (L. Li *et al.* 2018; Wu *et al.* 2019).

1.14.2 Titration based monitoring

Monitoring systems based on titration work by quantifying a substance of interest by determining the amount of a given reagent it reacts with (Spanjers and van Lier 2006). These methods can be used to measure parameters such as total VFA concentration, alkalinity and total ammonium (Ibarra-Esparza *et al.* 2025). For example the FOS/TAC method is a two point titration widely used in industry to estimate buffering capacity and total VFA concentrations, giving plant operators an idea of possible under- or overfeeding which could precede destabilisation (Liu *et al.* 2025). While titration based methods are relatively cheap and easy to operate (Ibarra-Esparza *et al.* 2025), disadvantages include

the fact that these techniques provide limited information when used in isolation, and can consume large amounts of chemicals (Wu *et al.* 2019).

1.14.3 Spectroscopic monitoring

Spectroscopy can be used to measure the concentration of components within a sample by detecting the absorbance, transmission and reflection of light (Wu *et al.* 2019). Ultraviolet/visible, infrared and fluorescence based methods have been used to measure a number of AD parameters including alkalinity, hydrogen sulphide production and VFAs (Mesquita *et al.* 2017). Near infrared (NIR) spectroscopy has been of particular interest because it allows non-destructive measurements to be performed rapidly (L. Li *et al.* 2018; Peng *et al.* 2022), but key drawbacks include complex calibration requirements and high equipment costs (Wu *et al.* 2019).

1.14.4 Chromatographic monitoring

Chromatographic techniques work through the principle of separating substances based on differences in their affinity between a mobile and stationary phase (Spanjers and van Lier 2006), and can be used to measure components in both the liquid and gas phases of AD (Ibarra-Esparza *et al.* 2025). For example gas chromatography with flame ionisation detection (GC-FID) is a well established approach that can be used to measure different VFAs separately in AD (Wu *et al.* 2019). A key disadvantage of chromatographic methods is that they generally require specialist instrumentation and skilled operators, making them expensive (Wu *et al.* 2019; Cruz *et al.* 2021).

1.14.5 Molecular biology techniques

Since the AD process is ultimately dependent on the microbiological community that underpins it, another possibility is to use molecular biology techniques to monitor this community directly (Venkiteshwaran *et al.* 2015; Amha *et al.* 2018; L. Li *et al.* 2018). Marker based approaches like 16S sequencing can provide an idea of community composition (Amha *et al.* 2018). Metagenomic sequencing approaches can be used to indiscriminately sequence DNA isolated from a community, giving a more detailed picture of its genetic capabilities (Vanwonterghem *et al.* 2014b), but conventional methods do not distinguish between DNA isolated from live, metabolically active and dead/inactive organisms (Hatzenpichler *et al.* 2016; Yap *et al.* 2022). Methods like biorthogonal non-

canonical amino acid tagging (BONCAT) overcome this issue by enabling only cells actively synthesising protein to be separated and sequenced (Hatzenpichler *et al.* 2016; Amekan 2021). The system works by growing microbial cultures in the presence of a synthetic methionine analogue which is incorporated into newly synthesised proteins, which are then labelled through a click-chemistry based reaction with a fluorescent dye enabling visualisation and fluorescence activated cell sorting (Hatzenpichler *et al.* 2016). In an alternative approach, incorporation of the methionine analogue into external membrane proteins has been used as an anchor for placement of a biotin tag which was subsequently used to physically extract only cells actively synthesising protein (Amekan 2021).

Transcriptomics, proteomics and metabolomics can be used to give a more detailed picture of the activity and outputs of the AD community than could be achieved with DNA based methods alone (Amha *et al.* 2018). For example, an early AD transcriptomics study identified that archaea were highly transcriptionally active (Zakrzewski *et al.* 2012) and a similar proteomics study found that while archaea comprised only around 4% of the community in the study system, proteins originating from them composed up to 30% of the total, suggesting that while they comprised only a small proportion of the community, the archaea were responsible for a considerable degree of its activity (Hanreich *et al.* 2013). A key limitation of 'omics based methods for monitoring purposes is that they are generally not suitable for real time monitoring and generate large amounts of data which requires skilled personnel to process (Innard and Chong 2022; Babu and Snyder 2023; Jin *et al.* 2025).

1.15 AD monitoring deployment

While the utility of some systems suitable for real-time monitoring of predictive parameters of stability such as NIR methods and some chromatographic procedures have been demonstrated, they are generally prohibitively expensive in an industrial setting (Wu *et al.* 2019). Additionally, many approaches have primarily been tested at laboratory rather than full scale (Cruz *et al.* 2021; Ibarra-Esparza *et al.* 2025). As a result, the AD monitoring methods which are *actually* widely used in industry are generally limited to simple physicochemical measurements such as pH or pressure, which have limited utility as early warning systems for possible process destabilisation (Wu *et al.* 2019). There

would therefore be great potential benefit in the development of faster, cheaper and more accessible methods for monitoring AD communities.

1.16 Engineered whole cell biosensors for AD community monitoring

Another option would be to use engineered biological approaches such as whole cell biosensors for AD monitoring. Whole cell biosensors are whole organisms which have been genetically "rewired" to produce a detectable signal in response to a chosen stimulus (Figure 1.1), and can be used, for instance, to report on environmental conditions such as the presence/absence of certain chemicals (Del Valle *et al.* 2020a). Biosensors exhibit some potential advantages over existing AD monitoring approaches in that (once developed) they could be cheap, easy to use, and report on the biologically relevant fraction of analytes (Renella and Giagnoni 2016; Innard and Chong 2022).

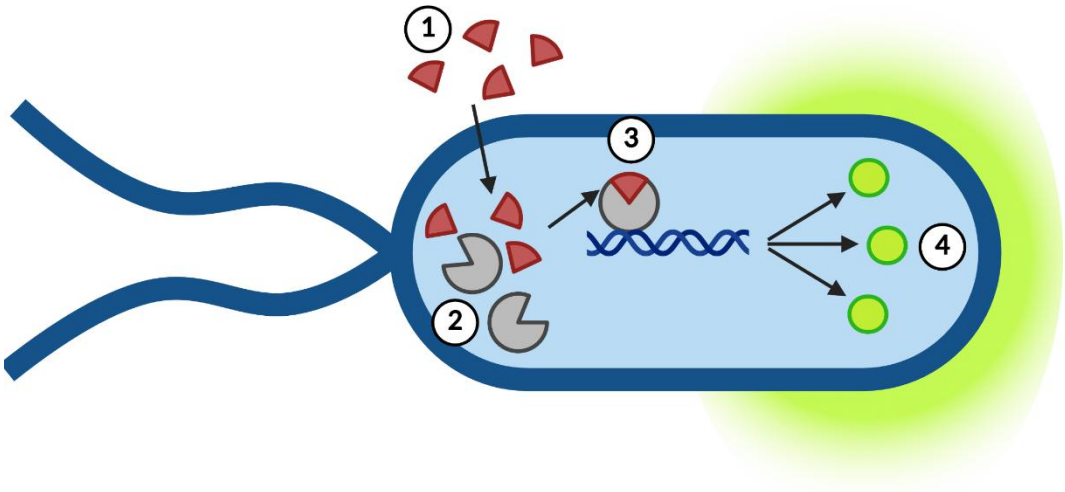


Figure 1.1: Core components of a whole cell microbial biosensor.
1) A chemical of interest (red triangle) the biosensor has been engineered to detect (e.g. some factor predictive of digester function) enters the cell. **2)** The chemical binds to a regulatory protein (grey circle). **3)** The chemical/regulatory protein complex initiates expression of a reporter gene (green circles). **4)** Once expressed, the reporter gene makes the cell easily detectable (fluoresces green). Adapted from (Innard and Chong 2022).

1.17 Existing whole cell biosensors

Engineered microbial whole cell biosensors have been developed for use in a variety of situations, and the field has progressed rapidly in recent years (Moraskie *et al.* 2021). One

area which has received considerable attention is environmental monitoring, particularly for xenobiotics such as heavy metals (Ali *et al.* 2021). For example, an early study described an *Escherichia coli* (*E. coli)* based biosensor engineered to detect arsenic contamination in drinking water as a potential alternative to established chemical testing methods (Stocker *et al.* 2003). Arsenic biosensors have undergone further development since, for instance the sensitivity of a similar system was increased by adding a genetically engineered positive feedback loop to amplify the fluorescent reporter gene output (Jia *et al.* 2019). Another notable deployment of microbial whole cell biosensors has been in medical diagnostics (Moraskie *et al.* 2021), for example *E. coli* was recently engineered to detect a virulence factor unique to *Pseudomonas aeruginosa*, a highly problematic bacterial pathogen associated with chronic infections in humans (Yue *et al.* 2024).

1.18 Reporter genes

The majority of whole cell biosensor studies (Ali *et al.* 2021; Moraskie *et al.* 2021), including the specific examples discussed above (Stocker *et al.* 2003; Jia *et al.* 2019; Yue *et al.* 2024), use reporters based on visual outputs such as fluorescence and luminescence. This is problematic in AD as existing well developed visual reporter systems like GFP do not work in this setting as they require oxygen (Streett *et al.* 2019). The currently available reporter gene toolbox for anaerobic environments is comparatively limited (Chia *et al.* 2019). Additionally, systems based on visual signals are difficult to use in opaque environments such as the contents of anaerobic digesters (Del Valle *et al.* 2020a). Since reporters are central components of biosensors (van der Meer and Belkin 2010), this presents a major barrier to their deployment in AD (Innard and Chong 2022).

1.19 Whole cell volatile biosensors

Volatile based reporter systems have recently been developed for use in opaque environments such as soil and sediments, and could be more practical than visual reporters for use in AD (Cheng *et al.* 2016). In these systems an enzyme which generates a volatile is expressed in reporter bacteria, and the gas produced is then detected in the system using an appropriate method such as gas chromatography-mass spectrometry (GC-MS) (Cheng *et al.* 2016; Innard and Chong 2022). One obvious application would be in AD biosensors: by expressing a volatile reporter system in conjunction with a suitable inducible promoter, the presence/absence of certain metabolites which are diagnostic of

AD community health could be detected non-invasively by the presence of the reporter volatile in the AD gas outlet (Figure 1.2).

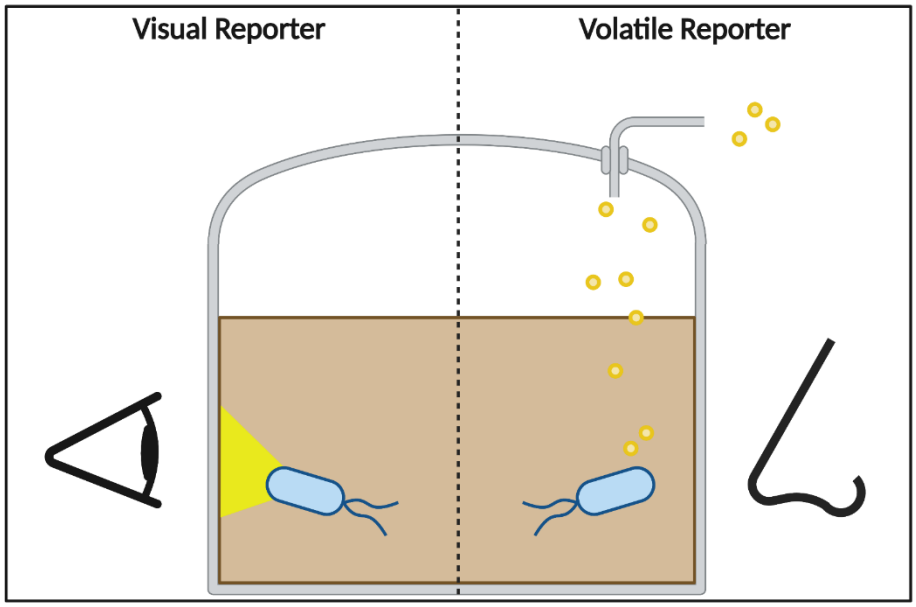


Figure 1.2: Potential advantages of volatile over visual reporters in AD. The visual signal produced by an engineered anaerobe (left) is difficult to detect within the anaerobic digester because the contents are non-transparent. The signal also cannot pass through the digester wall. An anaerobe engineered to code for a volatile producing enzyme (right) generates a gas (orange circles) which is otherwise absent within the digester. The volatile diffuses freely into the gas phase and is collected non-disruptively though the outlet pipe. Sampled effluent containing this gas indicates presence of the engineered anaerobe.

1.20 Existing whole cell volatile reporter systems

Genetically engineered volatile reporting is a relatively new technique which at the outset of this project had only been demonstrated in a handful of studies (Cheng *et al.* 2016; Cheng *et al.* 2018; Del Valle *et al.* 2020b; Fulk *et al.* 2020; Liu *et al.* 2020). In one study, engineered bacterial biosensors produced a methyl bromide forming enzyme in response to an inducer, generating methyl bromide gas proportional to the concentration of inducer in the medium. A constitutively expressed ethylene forming enzyme (EFE) gene was introduced into the same bacteria, resulting in the continual production of ethylene proportional to cell number. This system allowed quantitative reporting of inducer levels since the methyl bromide signal (produced only in response to the inducer) could be normalised to ethylene (produced constantly by all biosensor cells) (Cheng *et al.* 2018;

Innard and Chong 2022). If quantitative volatile reporter systems like this could be repurposed to work anaerobically, they could be highly useful in AD.

1.21 The importance of the microbial chassis

The chassis organism is the biological host into which genetic tools are placed, providing the cellular machinery necessary to allow those components to function as intended (Adams 2016). This means that the functioning of an engineered microbial tool such as a biosensor is not just reliant on the engineered genetic components working, but also on the survival of the host organism itself, since dead biosensors cannot perform their function (Close *et al.* 2009).

1.22 Existing chassis organisms

Widely used microbial chassis are generally model organisms with available genetic tools (Calero and Nikel 2019; Riley *et al.* 2019). These microorganisms have been artificially selected to be fitter in the environment they are reared in by humans (I.e. the lab), at the cost of fitness in the environment their ancestors were originally isolated from. They may additionally be rationally modified to make them easier to handle. For example, deletion of natural restriction modification (RM) systems which degrade foreign DNA can increase transformation efficiency (Foulquier *et al.* 2019), and the removal of "unnecessary" genes through genome reduction can be used to simplify the host, conserving cellular resources for the engineered function (Calero and Nikel 2019).

While useful in the lab, these changes come with the potential caveat of reducing chassis robustness in the natural environments from which their ancestors were originally isolated. For example, while deleting RM systems may make a bacterium easier to transform, these systems normally serve to protect the host from foreign DNA (Dimitriu *et al.* 2020), rendering the modified chassis less protected from harmful agents such as viruses. Similarly, while deleting genes deemed unnecessary in laboratory conditions may increase cellular productivity by channelling more resources towards growth, the lost genes may have important functions in other environmental contexts (Calero and Nikel 2019), leading to reduced fitness outside of the lab.

It follows that these domestic organisms would not be expected to survive as well in a "wild"/more natural setting due to the fitness cost of these artificially selected features (Steensels *et al.* 2019). This is unlikely to cause issues in an environment which can be easily tailored to suit the host: such as the laboratory, or even some industrial scenarios (where the organism can be cultivated in pure culture). However, these features are likely to be far more limiting in settings where conditions cannot be easily controlled (Calero and Nikel 2019; Steensels *et al.* 2019), as is the case in full scale anaerobic digesters (Puig-Castellví *et al.* 2022). The development of genetically tractable chassis organisms specifically for use in the AD environment could therefore be very useful.

1.23 AD community manipulation

As detailed above, AD process monitoring can aid the stability of digesters and their existing communities by potentially enabling the use of timely preemptive action. Another option is to manipulate the AD microbial community itself at the biological/organismal level with the objective of achieving a desired outcome, such as a more stable process or increased biogas yield (Vanwonterghem *et al.* 2014b; Carballa *et al.* 2015; Venkiteshwaran *et al.* 2015; De Vrieze and Verstraete 2016).

It is important to clarify what is meant here by microbial community manipulation, since several different approaches could potentially be described in this way. For example, the choice of operational parameters such as temperature, feedstock, loading rate, or supplementation with trace elements are known to be able to influence the microbial community (Carballa *et al.* 2015; Ferguson *et al.* 2016; DelaVega-Quintero *et al.* 2025). Reactive modifications such reducing feeding rate in response to digester acidification are used routinely with the aim of allowing the methanogenic portion of the community (and therefore the process as a whole) to recover (Nkuna *et al.* 2022; Shin *et al.* 2022). Alternative digester configurations such as two stage AD can enable separation of different steps of the AD process such that more optimal conditions can be provided to different subsets of the community (Rabii *et al.* 2019). While all of these interventions could reasonably be described as forms of microbial community manipulation, for the purposes of this thesis what is meant by this term is deliberate management of the composition and/or function of the community itself at the organism level in an attempt to achieve a desired outcome. In addition to improving process stability, this manipulation

could also potentially allow improved or novel functionality to be imparted on AD systems, such as improved biogas yields, the production of alternative high value end products besides biogas, or the degradation of unwanted products like microplastics (De Vrieze and Verstraete 2016; Harirchi *et al.* 2022; Tang 2023). Different approaches which could be used to achieve these are described below.

1.23.1 Bioaugmentation

Bioaugmentation is the addition of specific microorganisms with desirable features to existing digesters with the goal of improving performance in some way, and it has been the topic of considerable interest in the AD field (Chen *et al.* 2025; Alipoursarbani *et al.* 2026). A key approach is to augment the community with the aim of improving methane generation, for example in one study a pure culture of a lipid degrading organism *Clostridium lundense* was used to enhance AD of a lipid rich feedstock, resulting in an increased rate of methane production (Cirne *et al.* 2006). Another strategy is to add microorganisms to a struggling/destabilized system with the objective of aiding its recovery, for example a mixed culture of microorganisms isolated based on its ability to degrade the VFA propionate was used to augment experimentally overloaded systems, resulting in recovery times 25 days faster than controls (Tale *et al.* 2011). A notable advantage of bioaugmentation is that it could in principle be integrated into the existing treatment process (Tang 2023). One possible limitation is that introduced organisms may not successfully establish within the existing community, due to factors like competition with existing community members or difficulty adjusting to the new abiotic conditions (Giangeri *et al.* 2024). Another consideration is that the attempted bioaugmentation may have unpredictable outcomes (De Vrieze and Verstraete 2016), as evidenced by examples where bioaugmentation has not had the intended impact on AD performance (Westerholm *et al.* 2012; Fotidis *et al.* 2013; Palù *et al.* 2022).

1.23.2 Synthetic mixed communities

The potential of synthetic mixed communities (SynComs), which are defined as an artificially assembled group of microbial species (Nikoloudaki *et al.* 2024), has been explored in areas such as the human gut (Cheng *et al.* 2022) and agriculture (Gonçalves *et al.* 2023; Northen *et al.* 2024), and their potential future applications in AD were recently reviewed (Jourdain and Gu 2025). The general aim of a SynCom is to recreate the natural

functionality of a mixed microbial system while keeping it simple enough to be more easily controlled (Northen *et al.* 2024; Jourdain and Gu 2025). SynComs can be created by top-down methods, which involve enrichment of certain community members from an existing natural community, or bottom-up, whereby a synthetic community is assembled from scratch (Lawson *et al.* 2019). The field is still at a nascent stage in AD, with the majority of studies focused on top-down approaches (Jourdain and Gu 2025). Some recent studies employing small synthetic methanogenic communities composed of up to three (Chen *et al.* 2019) or four (D. Wang *et al.* 2023; Wang *et al.* 2025) different organisms demonstrate progress in bottom-up approaches. SynComs have progressed further in other areas, for example a bottom-up synthetic consortium of 100 different strains of human gut microbes has been produced (Perez *et al.* 2021), while SynComs containing mixtures of rhizosphere isolated bacteria and fungi were assembled to study their influence on susceptibility of tomato plants against Fusarium wilt disease (Zhou *et al.* 2022). One potential strength of SynComs as a microbial manipulation strategy is that they could theoretically be designed to have specific properties or achieve distinct goals (Lawson *et al.* 2019). It is also worth noting that they could also be used as the basis for bioaugmentation of existing systems (Lawson *et al.* 2019). A key outstanding challenge is maintaining the stability of a SynCom when deployed in the field, where a variety of abiotic and biotic factors may cause members of the community to be lost (Zhang *et al.* 2026)

1.23.3 Use of phage

Viruses that infect bacteria and archaea (phage) cause host cell lysis as part of their life cycles (Dar *et al.* 2025). Phage are ubiquitous in anaerobic digesters (Calusinska *et al.* 2016) and are thought to exert considerable impacts over microbial community composition and dynamics, either directly through cell lysis or indirectly through the release and subsequent utilization of cell components by other community members in a process called viral shunt (Zhang *et al.* 2017). One proposed use of viruses in AD is for the selective destruction of unwanted microbial community members such as pathogens or those involved in foam formation (Withey *et al.* 2005). Despite their important role in AD, research into these viruses has been limited compared with their bacterial and archaeal counterparts (Dar *et al.* 2025). Use of phage is more mature in other fields, for example phage cocktails have been applied to protect livestock animals and as preservatives in the

food sector, where commercial options are available (Cristobal-Cueto *et al.* 2021). Key benefits of the use of phage to manipulate microbial communities include that they can be highly specific to certain hosts (Thi Trang Nhung *et al.* 2025), and the fact that their reliance on their targets to reproduce means they are self-limiting (Gholizadeh *et al.* 2024). Current potential hurdles to their deployment include the need to generate large amounts of phage for water treatment, and the fact that microbial communities (i.e. their targets) may vary substantially between wastewater systems (Jassim *et al.* 2016).

1.23.4 Microbial genetic modification

Genetic engineering involves changing the genetic code of microbes in order to alter their phenotype (P. Yang *et al.* 2024). This approach is used to generate whole cell microbial biosensors for monitoring (see section 1.17), but can also be used for manipulating or enhancing community function (Bai *et al.* 2025). There has been increasing interest in modifying AD communities to produce other valuable end products besides methane such as VFAs or hydrogen (Khan *et al.* 2016). Genetic modification will be important to achieving this, particularly in cases where the desired trait is novel (Yan and Fong 2017; Cao *et al.* 2019; Liu *et al.* 2022; Elmore *et al.* 2023) such as the ability to synthesise a desired substance that does not normally exist in AD. For example, two species of *Methanosarcina* were engineered to produce isoprene, a valuable chemical precursor used in various consumer goods that they would not normally be able to produce (Aldridge *et al.* 2021).

AD presents an appealing platform for bioproduction because it does not require the addition of oxygen, which can be expensive (Kulakowski *et al.* 2023). This absence of oxygen potentially makes it even more attractive for certain chemicals. For example, ethylene is a valuable organic compound which is used widely as a precursor in the chemical industry in products such as plastics and surfactants (Fan *et al.* 2013). Aerobic pathways for ethylene biosynthesis are known (North *et al.* 2017) but a potential issue is that ethylene can be explosive in the presence of oxygen (Suga 1958). A biological route for anaerobic ethylene synthesis in some phototrophic bacteria was recently discovered (North *et al.* 2017). If this pathway could be engineered into a bacterial host suitable for use in AD it could provide a more economically viable route for ethylene biosynthesis.

One of the main advantages of genetic engineering for microbial manipulation is that it allows precise modifications to be made in a user defined manner (Jeong 2025). A key limitation is that developing genetic tools in host organisms to enable manipulations requires considerable time and effort (Sheth *et al.* 2016), particularly in the case of methanogens which have historically been more challenging to genetically manipulate, although recent developments in models like *Methanosarcina acetivorans* demonstrate progress in this area (Carr and Buan 2022; Myers and Dykstra 2024).

1.24 Aims

The overall aim of this project is to investigate biological approaches for monitoring and manipulating microbial communities in AD. As detailed above there are multiple different research areas which could be explored to achieve this, such as improving reporter genes for use in AD, developing chassis organisms specifically for the digester environment and working towards building defined bottom-up AD SynComs. Different approaches could also complement each other, for example chassis organisms are important for the development of both biosensors and as hosts for the engineered synthesis of valuable chemicals. Isolation of AD endogenous organisms could be used to generate SynComs but could also be used as the basis for the development of AD chassis organisms. Due to this potential synergy, the approach taken to achieving the overall aims of this project was to investigate several relevant research avenues concurrently:

1. To investigate an alternative to currently limited AD monitoring systems and reporter genes, the aim of Chapter 3 was to attempt to repurpose an existing whole cell volatile bioreporter for use in AD. The exogenous expression of a recently discovered anaerobic ethylene synthesis pathway was also investigated due to its potential utility as both a volatile reporter gene and as a valuable chemical precursor.
2. To demonstrate possible routes for obtaining chassis organisms more suited to the AD environment, two different approaches for obtaining or generating these based on selective enrichment of specific AD endogenous organisms or modification of laboratory organisms to be fitter in AD relevant conditions were investigated in Chapter 4.
3. The primary objective of Chapter 5 was to address the current gap in AD bottom-up SynComs by developing a high throughput strategy for obtaining and

cataloging organisms from AD samples with the long term goal of using these to generate defined AD communities.

Chapter 2: Methods

2.1 Molecular work

2.1.1 Construct design

Constructs were designed and assemblies simulated using the Benchling software (www.Benchling.com).

2.1.2 Codon optimization

Codons were optimized using the OPTIMIZER program (Puigbò *et al.* 2007). Sequences were input as DNA. Setting were kept on default except for the following: data from the HEG-DB (Escherichia_coli_str.K-12_substr._MG1655). Reference set: codon usage (HEG). Genetic code: 11: Eubacterial. Method used: guided random.

2.1.3 Gene synthesis

Designed genes were ordered as gBlocks (IDT). Lyophilized material was resuspended in TE buffer (10mM Tris, 0.1mM EDTA, pH 8) to the concentrations advised by the manufacturer.

2.1.4 Plasmid material

pHC01 (Cheng *et al.* 2018) and pEMF051 were a gifts from Jonathan Silberg. The following plasmids were obtained from the Addgene repository: pETM6_Ptrc_BsMdh_BmHps_BmPhi_Pthl_CaThl_CaCtfAB_CaAdc (Bennett *et al.* 2018) (Plasmid #129105), pJBEI_6409 (Alonso-Gutierrez *et al.* 2013) (Plasmid #47048), pBA2kIKmA2 (Lindberg *et al.* 2010) (Plasmid #39213). The Start-Stop Golden Gate kit plasmids (Taylor *et al.* 2018) were obtained from Addgene (Kit #1000000152).

2.1.5 Bacterial strains

E. coli strain MG1655λatt-kmgfp (Da Re *et al.* 2007) was a gift from Marjan van der Woude. *E. coli* strain MG1655_rLP5(nth-ydgR) (Urtecho *et al.* 2019) was obtained from Addgene (Bacterial strain #110245). *E. coli* strain ATCC 25922 was obtained from the UK Health and Security Agency (catalogue number NCTC 12241). *Bacillus subtilis* was obtained from Blades Biological Ltd (LZB 025).

2.1.6 Protein alignments

Protein alignments were performed using the Clustal Omega Multiple Sequence Alignment software (Madeira *et al.* 2024).

2.1.7 Re-dissolving plasmid material shipped on filter paper

Under sterile conditions, plasmid spots were excised from filter paper using a clean scalpel and fully immersed in 30 µl of pH 8 TE buffer (1 mM EDTA, 10 mM Tris-HCl) in an Eppendorf tube, then agitated by pipetting. Following a 10 min room temperature incubation, 2 µl of the resulting plasmid DNA (pDNA) sample was immediately transformed into chemically competent *E. coli*, and the remainder stored at -20 ^{o}C.

2.1.8 Plasmid sequencing

Plasmids were fully sequenced using the plasmidsaurus service (www.plasmidsaurus.com) as advised by the manufacturer.

2.1.9 Media preparation

2.1.9.1 LB broth

For 200 mL, 2 g of tryptone, 2 g NaCl and 1 g of yeast extract (LP0021 (Oxoid)) were added to 200 mL of dH_2O in a 250 mL Duran bottle or premixed power (LB Broth, Miller Modification (PhytoTech Labs)) was diluted as per the manufacturer's instructions. Media was shaken vigorously until all components had dissolved, then autoclaved at 121 ^{o}C, 15 psi for 20 min.

2.1.9.2 Nutrient broth

This was either made by adding 0.2 g of beef extract powder (Sigma-Aldrich), 0.4 g yeast extract (LP0021 (Oxoid)), 1 g trypticase peptone (BD) and 1 g sodium chloride to 200 mL of dH_2O in a 250 mL Duran bottle or premixed power (CM0001 (Oxoid)) was diluted as per the manufacturer's instructions. Media was shaken vigorously until all components had dissolved, then autoclaved at 121 ^{o}C, 15 psi for 20 min.

2.1.9.3 Premixed broth

The following broths were prepared as per the manufacturer's instructions: Brain Heart Infusion Broth (CM1135B, (Oxoid)), Reinforced Clostridial Medium (CM0149B (Oxoid)),

Anaerobe Basal Broth (CM0957 (Oxoid)). Broths were shaken then autoclaved at 121 ^{o}C, 15 psi for 20 min. Broths used for anaerobic culturing were transferred to the anaerobic chamber at least 24 hours in advanced with the caps loosened to allow equilibration.

2.1.9.4 IPTG

100 mM stock solution was prepared by dissolving 0.238 g of IPTG powder in 10 mL of sterile water. 1 mL aliquots were filter sterilized and sored at -20 ^{o}C.

2.1.9.5 X-gal

For 20 mg/mL stock solution, 0.2 g of X-gal was added to 10 mL of dimethylformamide, and vortexed until dissolved. 1 mL aliquots were stored in Eppendorfs wrapped in foil at -21 ^{o}C.

2.1.9.6 Antibiotics

Stock solutions were prepared at the following concentrations: Kanamycin 50 mg/mL, chloramphenicol 25 mg/mL, ampicillin 100 mg/mL, carbenicillin 100 mg/mL, tetracycline 10 mg/mL. Stocks were made by dissolving the appropriate amount of antibiotic powder in 20 mL of dH_2O (or ethanol for chloramphenicol) and filter sterilising through a prewet 0.2 μm filter into 1 mL aliquots, and stored at -20 ^{o}C.

2.1.9.7 LB agar

LB agar medium was prepared by adding 3 g of agar powder to 200 mL of LB medium (prepared as above), or using premixed power (LB-Agar Miller (Formedium)) diluted as per the manufacturer's instructions. Media was shaken vigorously then autoclaved at 121 ^{o}C, 15 psi for 20 min.

2.1.9.8 Schaedler Agar

1 L of medium contained 5. 67 g of casein hydrolysate (Millipore), 5 g of Bacto™ Peptone (gibco), 5 g of yeast extract (LP0021 (Oxoid)), 0.83 g of dipotassium hydrogen phosphate (fluorochem), 1.67 g of sodium chloride, 5.83 g of D-(+)-Glucose (Sigma-Aldrich), 0.4 g of L-cysteine (Sigma-Aldrich), 0.01 g of (porcine) haemin (Thermo Scientific) and 15 g of agar (Formedium). Haemin was first dissolved in 1 M NaOH. pH was adjusted to approximately

7.5 with 10 M KOH, autoclaved at 121 ^{o}C, 15 psi for 20 min and held in a water bath at 56 ^{o}C prior to plate pouring.

2.1.9.9 Premixed solid media

The following agar media were prepared as per the manufacturer's instructions: Brain Heart Infusion Agar (CM1136B (Oxoid)), Reinforced Clostridial Agar (CM0151B (Oxoid)), Anaerobe Basal Agar (CM0972B (Oxoid)), Nutrient Agar (CM0003, Oxoid)). Media was shaken vigorously and autoclaved at 121 ^{o}C, 15 psi for 20 min.

2.1.9.10 Agar plates

Solid media was melted in a microwave or boiling water bath and then allowed to cool to around 50 ^{o}C. Antibiotics or supplementary reagents were added if necessary to the following final concentrations: kanamycin 50 µg/mL, chloramphenicol 25 µg/mL, ampicillin 100 µg/mL carbenicillin 100 µg/mL, tetracycline 10 µg/mL, X-gal 200 µg/mL, IPTG 1 mM. Media was swirled, poured into 90 mm petri dishes in a laminar flow hood (20 – 25 mL per dish). Dishes were left open for approximately 1 hour to set, then lidded and stored at 4 ^{o}C. Plates used for anaerobic culturing were transferred into the anaerobic chamber at least 24 hours before use.

2.1.9.11 Premixed diluents

Quarter strength Ringer's solution (Ringer's solution tablets (quarter strength) (VWR Chemicals)) Maximum Recovery Diluent (MRD) (CM0733, Oxoid)), and Phosphate Buffered Saline (PBS) (BR0014G (Oxoid)) were prepared as advised by the manufacturers.

2.1.9.12 Resazurin Solution

1 g/L resazurin solution was made by dissolving 100 mg of resazurin powder (Sigma) in 100 mL of dH_2O.

2.1.9.13 Anaerobic LB broth supplemented with resazurin

200 mL of LB broth made up as explained in section 2.1.9.1 without autoclaving was mixed with 200 µL of resazurin solution (section 2.1.9.12) in a 1 L round bottom boiling flask. A rubber bung was inserted loosely into the neck of the flask with a gassing canula down the side of it into the vessel under a flow of 30 psi nitrogen such that displaced air could

escape. The flask was boiled with constant swirling over a Bunsen burner (blue flame) until the resazurin had changed from blue to pink to clear (indicated by the yellow colour of the LB medium). The amount of time this took was recorded for subsequent production of control media without resazurin. The cannula was then removed while pressing on the bung to close the vessel without oxygen ingress. The bung was taped in place and the vessel brought into the anaerobic chamber (COY) and allowed to cool. 5 mL aliquots were decanted into 150 mm anaerobic glass culture tubes which had been placed into the chamber at least 24 hours previously using a 25 mL serological pipette (Sarstedt) connected to Biohit Midi Plus pipette controller. Tubes were sealed with butyl rubber stoppers (also placed in the chamber 24 hours prior). Tubes were removed from the chamber, sealed with aluminium crimp camps (Fisher Scientific) and autoclaved at 121 ^{o}C, 15 psi for 20 min. Tubes without resazurin were prepared in the same way (including boiling for the same amount of time) but did not contain resazurin solution.

2.1.9.14 Anaerobic chloramphenicol solution

A 1.25 mg/mL chloramphenicol stock solution was made by adding 25 mg of chloramphenicol powder to 20 mL of 100% ethanol, dissolving and passing it through a prewet 0.2 µm filter membrane into a sterilised 100 mL glass vial (Wheaton) in the laminar flow hood. The vessel was sealed with a butyl rubber stopper and crimped with an aluminium crimp cap, sparged with nitrogen for 10 min, and stored at -20 ^{o}C.

2.1.10 *E. coli* culturing

E. coli was predominantly grown in LB medium or nutrient broth containing (where necessary) kanamycin to a final concentration of 50 µg/mL, chloramphenicol at 25 µg/mL ampicillin 100 µg/mL, carbenicillin 100 µg/mL or tetracycline 10 µg/mL. Inoculations were performed in a laminar flow hood by picking a colony from a plate or sample from a glycerol stock with a sterile pipette tip and dropping it into a 50 mL Falcon tube containing medium (normally 5 mL). Cultures were usually incubated overnight at 37 ^{o}C with 200 rpm shaking.

2.1.11 Glycerol stocks

Glycerol stocks were prepared by mixing 0.5 – 1 mL of bacterial culture with 0.5 – 1 mL 50% glycerol (1:1 ratio) in a laminar flow hood in 2 mL CRYO.S tubes (greiner). For

anaerobic cultures the liquid components were the same but stocks were prepared in 10 mL glass vials (Chromacol) and sealed with butyl rubber stoppers (which had been autoclaved wrapped in foil) in the anaerobic chamber. Vessels were closed with aluminium crimp caps. Stocks were stored at -80 ^{o}C.

2.1.12 Generating chemically competent cells (*E. coli*)

Chemically competent *E. coli* cells were prepared under sterile conditions predominantly using the Mix & Go! *E. coli* Transformation Kit and Buffer Set (Zymo Research) according to the manufacturer's instructions. *E. coli* cultures used for the protocol were grown in ZymoBroth™ at 25 ^{o}C with 200 rpm shaking. 100 µL aliquots of prepared competent cells were transferred to -80 ^{o}C for long term storage. For the calcium chloride method used for troubleshooting strain ATCC 25922 transformations the method was as described in (Chang *et al.* 2017) except the volumes were scaled down by 50% and competent cells were stored in 100 µL aliquots.

2.1.13 Plasmid DNA isolation

pDNA was isolated using the Nucleospin® Plasmid kit (MACHERY-NAGEL) from 5 mL overnight cultures according to the manufacturer's instructions under the "Isolation of high-copy plasmid DNA from *E. coli*" protocol. In instances where very low pDNA yields were obtained, this protocol was repeated except the elution buffer was heated to 70 ^{o}C, then the sample was incubated at 70 ^{o}C following addition of the elution buffer to the column for a further 2 min prior to centrifugation. In instances where this amended elution step also failed to yield sufficient pDNA for downstream applications, the "Isolation of low-copy plasmids, P1 constructs, or cosmids" protocol was used on 10 mL of overnight culture, and pDNA was eluted with pre-heated elution buffer as described above in a total volume of 30 µl. The QIAprep Spin Miniprep Kit (Qiagen) was also used as per the manufacturer's guidelines.

2.1.14 DNA quantification

DNA was quantified on a Nanodrop One^{c} (ThermoFisher Scientific) or a Qubit™ 4 Fluorimeter (Invitrogen) using the Qubit™ dsDNA BR Assay Kit and/or Qubit™ dsDNA HS Assay Kit as per the manufacturer's guidelines. pDNA was quantified only on the Nanodrop, genomic DNA (gDNA) was quantified by both methods (quantification using

High Sensitivity (HS) reagents was only performed if concentration using Broad Range (BR) reagents was too low for detection).

2.1.15 *E. coli* transformation

Chemically competent *E. coli* which had been prepared using the Mix & Go! *E. coli* Transformation Kit and Buffer Set (Zymo Research) were transformed using the protocol provided by the kit manufacturer with some modifications. 2 µL pDNA was added to a 100 µL aliquot of competent cells which had been thawed on ice, mixed gently by flicking and incubated for 10 min on ice. 400 µL of SOC medium (pre-warmed to 37 ^{o}C) was added and the sample incubated for one hour at 200 RPM shaking. 200 µL of transformation reaction was spread on an agar plate containing antibiotic appropriate to the plasmid, then incubated overnight at 37 ^{o}C. For the calcium chloride method used for troubleshooting strain ATCC 25922 the protocol was as described in (Chang *et al.* 2017) except 10 µL of pDNA was transformed into 100 µL of competent cells. For plating, transformed cells were centrifuged at 6000 xg for 1 min, supernatant removed to leave approximately 200 µl of liquid, the cells resuspended in this gently by pipetting and the whole suspension spread on a plate.

2.1.16 Diagnostic plasmid digestion

New England Biolabs (NEB) restriction enzymes were used to digest 1 µg of pDNA in 50 µL total reaction volumes according to the manufacturer's instructions on NEBcloner® version 1.9.0 (https://nebcloner.neb.com). Smaller reactions were scaled down appropriately.

2.1.17 Golden Gate assembly reactions

Golden Gate reactions using the Stop-Start kit were conducted according to the protocol in (Taylor *et al.* 2018).

2.1.18 Agarose gel electrophoresis

DNA was separated on gels containing agarose powder dissolved to a final concentration of 1% in TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA) and SYBR® Safe DNA Gel Stain (ThermoFischer Scientific) to a final concentration of approximately 1x. 6 µL of premixed Generuler 1 kb Ladder (ThermoFischer Scientific) was loaded to estimate band

sizes. Orange Loading Dye (6x) (ThermoFischer Scientific) was added to samples to a final concentration of 1x. Gels were run at 100 V until the loading dye had visibly migrated approximately three quarters of the way down the gel then imaged.

2.1.19 Genomic DNA extractions

2.1.19.1 From sludge

gDNA was extracted from sludge samples using the DNeasy® Powersoil® Pro Kit (Qiagen) according to the manufacturer's instructions with the following amendments. Approximately 2 mL of sludge was transferred to a 2 mL Eppendorf tube and pelleted by centrifugation at 16160 xg for 3 minutes. Supernatant was discarded and 250 mg of pelleted sludge was transferred to a fresh 1.5 mL Eppendorf tube with a wooden spatula. 800 µL of buffer CD1 was added, the tube vortexed and the sample was transferred to a PowerBead Pro tube. After step 14 in the manufacturer's protocol an additional wash was performed by adding 500 µL of freshly prepared 70% ethanol, centrifuging the tube at 15000 xg and discarding the flow-through. Samples were eluted in 50 µL of buffer C6.

2.1.19.2 From microbial cell cultures

DNA was extracted from pure or pooled microbial cell cultures using the DNeasy UltraClean Microbal Kit (Qiagen) with 1.8 mL of culture as per the manufacturer's instructions.

2.1.19.3 From platewashes

Bacterial colonies on plates were recovered as pellets as explained in (Alessi *et al.* 2018) except using 5 mL of PBS per plate. In cases where there were multiple replicate plates, washes were pooled prior to pelleting. For the T and S medium platewashes (section 2.1.25.2) an extra step in which a sterile disposable spreader was used to gently agitate the surface of the plate with the liquid on was used before pelleting.

2.1.19.4 For kit comparison for 8 organism test pool experiment

5 mL liquid cultures for the 8 selected organisms were generated by inoculating their respective glycerol stocks into the broth corresponding the original medium each had been isolated from (NA for organisms A, B, C, RCA for D, E, F and BHI for G, H) in 50 mL Falcon tubes. Cultures were incubated at 35 ^{o}C, 150 RPM shaking for 4 days. 1.8 mL of

each culture was transferred to separate 2 mL Eppendorf tubes then the remainder of all cultures were mixed together by decanting to create a pooled sample. 1.8 and 3.6 mL samples of the pooled culture were extracted using three different kits: the DNeasy UltraClean Microbal Kit (Qiagen) (referred to in the results section as “kit 1”), the DNeasy Powersoil Pro Kit (Qiagen) (“kit 2”) and the ZymoBIOTICS DNA Miniprep Kit (“kit 3”). For all kits for the 3.6 mL sample, half was first loaded, pelleted as advised, then the supernatant poured off and the remainder loaded on top and pelleted again. The DNeasy UltraClean Microbal Kit (Qiagen) was used otherwise as per the manufacturer’s guidelines. For the DNeasy Powersoil Pro Kit (Qiagen) 800 µL of solution CD1 was used to resuspend pelleted cells and the suspension was transferred to the PowerBead Pro tube. For the ZymoBIOTICS DNA Miniprep step 1, 750 µL of lysis solution was instead used to resuspend pelleted cells and the whole resuspension was transferred to the bashing bead lysis tube. Optional wash steps 12 and 13 were included. The vortex step for all kits was done in a Vortex-Genie® 2 on the highest speed setting (10). The 8 separate 1.8 mL culture samples aliquoted earlier were extracted using the DNeasy UltraClean Microbal Kit (Qiagen) as per the manufacturer’s guidelines.

2.1.19.5 For 96 organism test array pooling

Aerobic isolates and glycerol controls were inoculated into 5 mL of the respective medium they had originally been isolated from in 50 mL Falcon tubes. Anaerobic isolates and glycerol stocks were inoculated in the anaerobic chamber into glass test tubes containing 5 mL of their respective medium which had been placed in the chamber with loosened caps at least 24 hours in advance. Inoculated tubes were closed with butyl stoppers, removed from the chamber and crimped with aluminium caps. For both aerobic and anaerobic inoculations, one mock inoculation was performed for each of the three media types (NA, BHI and RCA) by agitating a sterile pipette tip in the media then processing the tubes as for the other inoculations. All cultures were incubated at 35 ^{o}C, 150 RPM shaking for 4 days. Liquid cultures were laid out in a 96 sample array (Figure 4.23) then pooled by rows and columns by pipetting 500 µL of each component culture into a falcon tube to generate 12 eight sample column and eight 12 sample row mixes (20 samples total). 1.8 mL of each pooled sample was extracted using the DNeasy UltraClean Microbal Kit (Qiagen) as per the manufacturer’s instructions.

2.1.20 Sludge sampling

All sludge used in this project was collected from a full-scale 3433 m^3 anaerobic digester at the Yorkshire Water Esholt sewage treatment works, which treats 760,000 population equivalent of wastewater sludge. Material was transported to the laboratory in a sealed 30 L clip top plastic drum (Mauser) and stored at 4 ^oC until processing, which was generally within 24 hours. Sludge at the site is thermally hydrolysed prior to AD. Pre- and post-THP material was collected just before or just after the thermal hydrolysis step then transported and processed as described for sludge.

2.1.21 Sludge plating

Sludge samples were homogenised with a metal augur, a subsample was collected in a 50 mL Falcon tube and used to create a series of 10-fold serial dilutions in MRD, PBS or quarter strength Ringer's solution. The number of dilutions varied by experiment but was always conservative to ensure an appropriate low number of colonies suitable for counting would be present. 100 or 200 µL of each dilution was plated in triplicate for each medium type used in the respective experiment. Control plates were prepared in the same way but only the respective dilatant was plated. All aerobic plating was performed in a laminar flow hood. Anaerobic plating was performed similarly, but plates, dilatants and reagents were transferred into the anaerobic chamber at least 24 hours prior to the work and the plating itself was done in the anaerobic chamber. Anaerobic plates were incubated at the required temperature in an anaerobic airtight vessel or in a Model 2000 Forced Air Incubator (COY Laboratory Products) inside an anaerobic chamber (COY Laboratory Products).

2.1.22 Colony counting and colony forming unit determination

Unless otherwise stated, plate counts and calculations were performed on the set of triplicate plates within a dilution series where colonies were in the range of 30-300. Plates were placed on top of an LED light box (Light4vision), the petri dish was overlaid with a glass cover and colonies were counted by eye using a marker pen to note previously counted colonies. Colony forming units (CFU) per mL of wet sample were calculated by multiplying the number of colonies counted on the plate by the reciprocal of the dilution factor then dividing this by the volume of sample plated in mL. In rare cases where colonies were recovered on control plates these counts were subtracted from their

corresponding test plates, apart from in cases where a control plate had clearly been compromised in some way (e.g. was visibly damaged) in which case it was omitted. Fluorescent colonies on plates were imaged using an iBright Fl1000 imaging system (Invitrogen), then colony counting and CFU calculations were performed using generated images as above.

2.1.23 Experimental evolution

2.1.23.1 Mixed strain plate comparisons

E. coli strains MG1655λatt-kmgfp (Da Re *et al.* 2007), MG1655_rLP5(nth-ydgR) (Urtecho *et al.* 2019) and MG1655 were streaked separately on LB agar plates and incubated overnight at 37 ^{o}C. Individual colonies from each plate were used to inoculate 5 mL of LB broth in 50 mL Falcon tubes and incubated overnight at 35 ^{o}C, 150 RPM shaking. On ice in the laminar flow hood cultures were adjusted to an OD_{600} of approximately 0.05 and the value recorded. The cultures were then mixed in the three pairwise combinations by adding 500 µL of each to a 2 mL Eppendorf tube. 10-fold serial dilutions of each mix were made in LB medium to 10^{-7}. 200 µL of all dilutions for each mix was plated in triplicate on LB agar plates then incubated overnight. Imaging, counting and CFU calculation were as explained in section 2.1.22 except the counts were adjusted to the exact starting OD_{600} value for each strain.

2.1.23.2 Resazurin impact experiment

E. coli strain MG1655_rLP5(nth-ydgR) (Urtecho *et al.* 2019) was streaked on LB agar supplemented with chloramphenicol and incubated overnight at 37 ^{o}C. Individual colonies were used to inoculate 5 mL of LB broth supplemented with chloramphenicol aerobically in 50 mL Falcon tubes, as well as an uninoculated control and incubated overnight at 35 ^{o}C, 150 RPM shaking. For all subsequent steps anaerobic LB broth with and without resazurin prepared as described in section 2.1.9.13 was used, liquid transfers were performed by using a 21G x 1½" Nr 2 Microlance™ (BD) hypodermic needle pierced through the butyl rubber stopper and OD_{600} measurements were performed using the Spectronic 20D+ spectrophotometer to maintain anaerobic conditions. The overnight culture was first added to an anaerobic LB tube without resazurin and OD_{600} adjusted to 0.05. The same amount of uninoculated control overnight culture was added to a second anaerobic LB tube without resazurin. 0.1 mL of anaerobic chloramphenicol solution (see

section 2.1.9.14) was added to 12 anaerobic LB broth tubes prepared with resazurin and 12 without resazurin. Six of each tube type was made aerobic by passing 3 60 mL syringes of air through them in a laminar flow hood using a 21G x 1½" Nr 2 Microlance™ (BD) hypodermic needle pierced through the rubber stopper and a second needle (not attached to the syringe) to allow headspace to escape. 0.1 mL of the OD_{600} 0.05 adjusted overnight culture was transferred into three of the aerobic tubes with and without resazurin creating treatments "inoculated + resazurin + O2" and "inoculated – resazurin +O2", and three of the anaerobic tubes with and without resazurin creating treatments "inoculated + resazurin - O2" and "inoculated – resazurin - O2". The uninoculated LB tube was used to prepare the four corresponding control treatments in triplicate in the same way. All tubes were inverted three times then OD_{600} measured to give t = 0 measurements. Tubes were incubated at 35 ^{o}C, 150 RPM shaking for 30 hours and OD_{600} was measured as explained above every hour for the first 16 hours then every 2 hours for the final 14.

2.1.23.3 Evolution passage

An OD_{600} ≈ 0.05 culture of *E. coli* strain MG1655_rLP5(nth-ydgR) (Urtecho *et al.* 2019) was prepared as explained in section 2.1.23.2, and used to inoculate eight anaerobic LB broth tubes with resazurin prepared as explained in 2.1.9.13, all of which had been supplemented with 0.1 mL of chloramphenicol solution (see section 2.1.9.14) and four of which had been made aerobic as explained in section 2.1.23.2 to create anaerobic populations 1-4 and aerobic (control) populations 5-8. OD_{600} readings were taken using a Spectronic 20D+ spectrophotometer then all populations were incubated at 35 ^{o}C, 150 RPM shaking for 14 days. Each day, the cultures were removed from the incubator, OD_{600} readings were taken, and 1 mL syringes with 21G x 1½" Nr 2 Microlance™ (BD) hypodermic needles were used to passage 0.1 mL of each culture to a fresh LB broth tube of the corresponding type prepared as explained above. The passaged tubes were measured at OD_{600} and returned to incubation. On days 7 and 14 overnight cultures were also used to prepare and store a separate glycerol stock for each population as explained in 2.1.11, except the *E. coli* culture was withdrawn using a syringe and needle as explained above.

2.1.23.4 Comparative growth experiment

E. coli strain MG1655_rLP5(nth-ydgR) (Urtecho *et al.* 2019) (ancestral) and glycerol stocks of the 8 evolved populations prepared after 14 days of passage were inoculated into LB

broth supplemented with 0.1 mL of chloramphenicol solution (see section 2.1.9.14). Control inoculations of *E. coli* MG1655 (prepared in the same way), and a mock inoculation containing no bacteria were also prepared. All cultures were incubated aerobically at 35 ^{o}C, 150 RPM shaking for 25.5 hours, then on ice all were diluted to OD_{600} ≈ 0.05 in anaerobic LB broth tubes as explained in section 2.1.23.2 (control equivalents had no growth so were prepared by transferring 0.1 mL of aerobic culture to the anaerobic vessel). 33 anaerobic LB broth tubes containing resazurin which had been prepared as explained in section 2.1.9.13 had 0.1 mL of anaerobic chloramphenicol solution (see section 2.1.9.14) added. On ice, each of the diluted cultures was used to inoculate these in triplicate by transferring 0.1 mL to each of three tubes using a 1 mL plastic syringe (Plastipak™ (BD)) and 21G x 1½" Nr 2 Microlance™ (BD) hypodermic needle in a laminar flow hood. OD_{600} readings for these cultures were taken (t = 0 measurements) on the Spectronic 20D+ spectrophotometer, and the cultures were incubated at 35 ^{o}C, 150 RPM shaking for 124 hours. Growth readings were taken as explained above every hour for the first 18 hours (excluding between hours one and four due to an issue with the spectrophotometer) and a final reading was taken after 124 hours.

2.1.24 Selected organism isolation on plates

2.1.24.1 *E. coli*

Selective plating for coliforms including *E. coli* was performed as described in Part B of (The Standing Committee of Analysts 2016) with the following amendments. Membrane Lactose Glucuronide Agar (MLGA) (CM1031 (Oxoid)) plates were prepared as advised by the manufacturer. 1 mL of sludge (see section 2.1.20) was serial diluted to 10^{-7} in MRD and plated as described in section 2.1.21. The 37 ^{o}C incubation step was performed for 44 hours.

2.1.24.2 *B. subtilis*

Selective plating for *Bacillus* was as described in (Environment Agency 2007) with the following amendments. Nutrient agar was CM0003 from Oxoid and prepared as advised except for the addition of bromothymol blue solution as described in (Environment Agency 2007). pH of the medium was adjusted from 7.17 to 7.22 with 1 M NaOH. 1 mL of sludge (section 2.1.20) was transferred to a 15 mL Falcon tube and heated in a water bath at 60 ^{o}C for 20 min, cooled in cold water, then serial diluted to 10^{-7} in quarter strength

Ringer's solution and plated as described in section 2.1.21. The 37 ^{o}C incubation step was performed for 48 hours.

2.1.24.3 *C. beijerinckii*

Selective plating for *C. beijerinckii* was as described in Part A of (The Standing Committee of Analysts 2021) with the following amendments. Tryptose sulphite cycloserine agar without egg yolk was instead Perfringens Agar Base (TSC & SFP) (CM0587 (Oxoid)) and prepared as advised except for the addition of 100 mg/mL D-cycloserine solution as described in (The Standing Committee of Analysts 2021). 1 mL of sludge (section 2.1.20) was transferred to a 15 mL Falcon tube and heated in a water bath at 60 ^{o}C for 15 min then transferred to the anaerobic chamber, serial diluted to 10^{-7} in MRD and plated anaerobically as described in section 2.1.21. The 37 ^{o}C incubation step was performed in the anaerobic chamber for 48 hours.

2.1.24.4 Colony selection and liquid culture outgrowth

Colonies from *E. coli* and *B. subtilis* plates were picked into 5 mL of nutrient broth (CM0001 (Oxoid)) in 50 mL Falcon tubes in a laminar flow hood. Colonies from *C. beijerinckii* plates were inoculated into 5 mL of Reinforced Clostridial Medium (CM0149B (Oxoid)) in glass test tubes in the anaerobic chamber, sealed with butyl rubber stoppers and crimped with aluminium crimp caps (all reagents and materials had been placed in chamber at least 24 hours earlier). All cultures were incubated at 35 ^{o}C, 150 RPM shaking overnight prior to gDNA extraction except for colony 3 from the *B. subtilis* plate (NAB 3) which was incubated for 7 days under these conditions before gDNA extraction.

2.1.25 Separately autoclaved media plating

2.1.25.1 R medium

For the first experiment R medium 1bis was prepared as described in (Dione *et al.* 2016) except haemin was not included in solution B, and instead dissolved separately in 1 mL of 1 M NaOH, mixed with 9 mL of dH_2O, stirred for 10 min, autoclaved at 121 ^{o}C, 15 psi for 20 min and mixed with the other components just before plate pouring. Schaedler agar was prepared as explained in section 2.1.9.8. Sludge sampling was as in section 2.1.20. Serial dilution down to 10^{-7} using PBS followed by plating in triplicate at all dilutions onto both R medium and SA was as explained in 2.1.21. Plates were incubated aerobically at

35 ^{o}C. The second experiment was carried out as explained above except the pH of solution C was adjusted to 7.6 with 1M NaOH and only 0.01g of haemin was added per L of medium (i.e. 10-fold less than the R medium recipe and the same as in standard Schaedler agar which also contains haemin (Dione *et al.* 2016)).

2.1.25.2 T and S medium

Both media were prepared as explained in (Kawasaki and Kamagata 2017) where they are referred to as "PT" and "PS" medium, respectively. The 10x phosphate buffer solution was adjusted to pH 6.8 using 1 M NaOH and 5 M KOH. Sludge sampling was as in section 2.1.20. Serial dilution down to 10^{-5} using MRD followed by plating 6 plates at all dilutions onto both T and S medium was as explained in 2.1.21. Three of each plate type at each dilution were incubated aerobically, and three of each anaerobically at 35 ^{o}C for 7 days.

2.1.26 Colony picker work

All work was conducted on a PIXL Dark (Singer Instruments) colony picker. A UV sterilisation cycle was run for 20 min after each use. Plates lids were always removed only once inside the instrument, and replaced before removal.

2.1.26.1 Colony picking and inoculation into liquid culture

The donor agar plate containing colonies (always on 90 mm petri dishes) was loaded into the corresponding source bay of the colony picker and the "Random Colony Picking" protocol was selected to identify colonies using default settings except selection algorithm set to "Background Subtraction – *B. subtilis*". If colony identification was poor (e.g. misidentification of colonies or many colonies missed) the protocol was cancelled and these steps repeated. If colony detection remained poor despite multiple attempts the run was cancelled again and the manual selection function was used to improve the identification. Colonies were picked onto liquid media in either a deep (v bottom polypropylene 96-well Costar Assay Block, 2 mL (Corning)) or shallow (flat bottom polystyrene 96-well Costar Cell Culture Plate (Corning)) 96 well plate with wells containing 1 mL or 200 µl (respectively) of broth loaded in the rectangular source bay using the appropriate default program dependant on the purpose of the experiment.

2.1.26.2 *E. coli* isolation and outgrowth test

E. coli strain MG1655 was inoculated into 5 mL of nutrient broth and incubated overnight at 35 ^{o}C, 150 rpm shaking. The culture was serial diluted in MRD to 10^{-6}, all dilutions were plated in triplicate on nutrient agar and incubated overnight at 35 ^{o}C. Resulting colonies were transferred using the colony picker into three shallow plates ((flat bottom polystyrene 96-well Costar Cell Culture Plate (Corning)) and three deep plates (v bottom polypropylene 96-well Costar Assay Block, 2 mL (Corning)) contained 200 µL (shallow) and 1 mL (deep) of sterile nutrient broth in each well in the pattern shown in Figure 5.9. Picking settings were as follows. Pinning: well base clearance 5% of well. Mixing: wet - on. radius: 75% of available. Vertical travel: 25% of available. All plates were sealed with clear polyester sealing film (starlab), lidded and secured with tape. Plates were incubated overnight at 35 ^{o}C. One of each type was incubated at three different speeds: Vortemp 56 incubator (Labnet International) at 1000 rpm (fast), a Thermo Scientific MaxQ 8000 Incubator at 300 RPM (slow) and a static incubator (Gallenkamp).

2.1.26.3 *E. coli* glycerol storage and revival tests

The statically incubated deep well plate from section 2.1.26.2 was replica plated into 3 shallow plates (flat bottom polystyrene 96-well Costar Cell Culture Plate (Corning)) in a laminar flow hood by transferring 100 µL of *E. coli* culture into each well which had been pre-filled with 100 µL of 50% glycerol solution using a multichannel pipette. Plates were sealed with clear polyester sealing film (starlab), lidded, and frozen at -80 ^{o}C and used as source plates to inoculate fresh shallow plates in triplicate using 3 different methods. For method one a frozen source plate was removed from -80 ^{o}C on dry ice. Using a multichannel pipette sterile tips were inserted into the frozen material in the first column of the source plate and used to transfer a small amount of frozen culture to a recipient plate which had been pre-filled with 200 µL of nutrient broth in all wells. This was repeated for the remainder of this plate and two further plates. For method two a source plate was thawed and the colony picker used to rearray the thawed material into 3 destination plates each with 200 µL of nutrient broth in all wells using the rearray function with pinning set to well base clearance: 5% of well. In method three the source plate was thawed and 20 µL of culture was transferred from the first column into the corresponding column of a destination plate containing 180 µL of nutrient broth in all wells. This was repeated for the remainder of this plate and 2 further plates. All plates were sealed with

clear polyester sealing film (starlab), lidded, and incubated static at 35 ^{o}C overnight. For the deep well test the statically incubated deep well plate from section 2.1.26.2 was frozen at -80 ^{o}C, thawed and the colony picker was used to rearray the contents into three deep well plates each containing 1 mL of nutrient broth in all wells using the rearray function with pinning set to well base clearance: 5% of well. Plates plate sealed, lidded and incubated 35 ^{o}C static for approximately 30 hours.

2.1.27 Ethylene sampling and analysis (photoionisation detector)

Ethylene was measured using a MiniPID2 11.7 eV photoionisation detector (Ion Science) held within a series of gastight plastic containers connected via an AI/O USB6001 DAC box (National Instruments) to a laptop using LabVIEW (version 21.0) software. 10 mL samples held within a 10 mL gastight syringe were injected into one of two 2-way Luer stopcocks connected to the gastight container via a 1 cm plastic tube. Immediately after sample injection both stopcocks were shut, and a reading was displayed on the laptop screen. The system was prepared for the next sample by reopening both Luer stopcocks and passing lab air from a 60 mL plastic syringe through the system until the reading stabilised back around the baseline value. Investigations were performed using neat 5000 ppm ethylene/nitrogen mix (BOC).

2.1.28 Sludge viability experiment

2.1.28.1 Sludge starving

Within 24 hours of arrival in the lab, a sludge sample obtained as explained in section 2.1.20 was homogenised with a metal augur and then approximately 1800 mL was transferred to a 2 L Duran bottle, which was made gastight by sealing with a GL 45 gastight closure and open screw cap (Duran). The headspace was sparged with nitrogen for 10 min. The vessel was then connected to a second bottle prepared in the same way except filled with 1800 mL of dH_2O via a 40 cm stretch of plastic tubing. One end of the tubing was connected to the water bottle via a Luer valve with a Luer syringe needle which was pierced through the gastight closure and immersed in the water. The other end of the tubing was connected to the sludge bottle via a Luer valve with a 21G x 1½" Nr 2 Microlance™ (BD) hypodermic needle pierced through the closure but not immersed in the liquid. The bottles were incubated at 35 ^{o}C and starved for 75 days, at which point no

further gas production was visually evident (i.e. no further air bubbles were seen emerging from the syringe needle immersed in water).

2.1.28.2 BMP test

The starved sludge (section 2.1.28.1) was disconnected from the gas relief system and brought into an anaerobic chamber, homogenised by swirling, and 34 approximately 27 mL samples were decanted into 50 mL falcon tubes. 17 tubes were transferred to -20 ^{o}C, while the other 17 were incubated at 35 ^{o}C. After 12 hours the frozen samples were thawed for 3 hours at 35 ^{o}C and then frozen and not-frozen samples were homogenised in sterile 500 mL Duran bottles. Cellulose was added as a feedstock in an inoculum: feed substrate volatile solid (VS) ratio of 2:1 as explained in (Kirby *et al.* 2020). In the anaerobic chamber 0.681 g of cellulose powder and 0.319 µL of distilled water was added to eight 100 mL serum bottles, and 1 mL of distilled water was added to another eight bottles. With regular swirling, 50 mL of previously frozen sludge was transferred into 4 of the serum bottles containing cellulose (fed treatment) and 4 without (unfed). The samples were mixed vigorously by swirling immediately after sludge addition. The process was repeated for the not-frozen sludge. All bottles were sealed with butyl rubber stoppers and aluminium crimp caps and then sparged with nitrogen for 10 min. The pressure of all bottes was equalised by first incubating at 35 ^{o}C for 6 hours and then withdrawing headspace. This was done through a 10 mL gastight glass syringe (SGE) connected to both a pressure transducer (Gems™ Sensors & Controllers) and 20 µm filter with a 21G x 1½" Nr 2 hypodermic needle via a three-way Luer stopcock. The needle was inserted through the butyl stopper of the bottle, pressure measured and headspace withdrawn until all bottles had a starting pressure reading corresponded to atmospheric pressure. All bottles were transferred to a 35 ^{o}C incubator. Every two to three days bottles were removed from the incubator in turn and pressure recorded as explained above. If pressure exceeded a threshold pre-determined to correspond to 6.5 mL of headspace, sufficient gas had been generated to withdraw a 5 mL sample which was done in a 10 mL gastight glass syringe as explained above then quantified for methane content (section 2.1.29). Further headspace gas was then removed until the pressure reached atmospheric levels, then this value and the total amount of gas withdrawn in mL was recorded.

2.1.28.3 Sludge storage

Immediately after the starved sludge sample was used to prepare the BMP test (section 2.1.28.2) the remainder was decanted in approximately 2 mL volumes into 320 15 mL Falcon tubes in the anaerobic chamber. The bottle containing the starved sludge sample was swirled in between aliquots. The tubes were sealed, withdrawn from the chamber and transferred to -20 ^{o}C for long term storage.

2.1.29 Methane quantification

Percentage methane content was measured using an infrared methane sensor (Dynament) held within a gastight plastic container connected via an AI/O USB6001 DAC box (National Instruments) to a laptop using LabVIEW (version 21.0) software. 5 mL samples held within a 10 mL gastight syringe were injected into one of two 2-way Luer stopcocks connected to the gastight container via a 1 cm plastic tube. Immediately after sample injection both stopcocks were shut, and a reading was recorded once stabilised. The system was prepared for the next sample by reopening both Luer stopcocks and passing lab air from a 60 mL plastic syringe through the system until the reading stabilised back around the baseline value. Calibrations were performed on each day of sampling using lab air and calibration gases containing 5, 20 and 60% methane (fti).

2.2 GC-MS sampling and analysis

2.2.1 Originally used GC-MS

2.2.1.1 Acetone plasmid growth experiment sampling

Glycerol stocks of *E. coli* MG1655 containing plasmid pETM6_Ptrc or no plasmid were streaked on LB agar plates containing ampicillin or no antibiotic, respectively and incubated at 37 ^{o}C for 18 hours. A single colony from each plate was then used to inoculate a 50 mL Falcon tube containing 5 mL LB with ampicillin for the plasmid containing strain or no antibiotic for the other, and incubated at 200 RPM, 37 ^{o}C for a further 18 hours. These cultures were then used to inoculate 700 mL of LB broth to a starting OD_{600} of 0.05. Ampicillin was added to the culture containing pETM6_Ptrc. The cultures were swirled then 200 mL of each was transferred into three separate sterile 2 L Duran bottles, which were sealed with bromobutyl rubber stoppers and GL-45 open topped screw caps to make them airtight. Vessels were incubated at 200 RPM, 37 ^{o}C and OD_{600} measurements were taken every half an hour until growth was stationary, at which

point headspace was sampled in 500 mL electropolished stainless steel canisters. All samples were analysed on a GC-MS less than 2 weeks after sampling.

2.2.1.2 Acetone/isoprene spiking experiment sampling

The desired dilutions were prepared by injecting a few drops of pure liquid standard (isoprene: Alfa Aesar, acetone: Sigma-Aldrich) into airtight 120 mL Wheaton vials closed with butyl stoppers and leaving them to stand for 1 hour for the volatile to equilibrate into the gas phase according to its vapour pressure. A dilution was then made by transferring the required amount of headspace gas to achieve a 2000 ppm concentration into a second stoppered vial with a gastight syringe. The final dilution was made by withdrawing 3 or 4 mL of 2000 ppm headspace into the airtight syringe, inserting it into an empty airtight 2 L Duran, expelling 1 ml, locking the syringe again and repeating for the other 2 L Durans. 2 L bottles were incubated for 10 hours at 200 RPM then sampled in 500 mL electropolished stainless steel canisters. Control bottles were prepared in the same way, except the initial Wheaton bottle contained no acetone standard, and a separate syringe was used. All samples were analysed on a GC-MS less than 2 weeks after sampling.

2.2.1.3 Volatile analysis

Volatile samples previously collected in 500 mL electropolished stainless steel canisters were condensed on a liquid nitrogen trap then passed to a Restek© Porabond Q column (30 m, 0.32 mm internal diameter, 0.5 µm thickness) inside a HP 5927 MSD mass spectrometer by immersing the trap in boiling water just after exposure to a helium flow. The GC was set to a flow rate of 4 mL/min for the first 2 min of the run, followed by 1 mL/min for the remainder. The oven temperature was set at 15 ^{o}C for the first 15 min, then increased by 2.5 ^{o}C/min up to a temperature of 103 ^{o}C and held for 5 min. The temperature was then increased by 4 ^{o}C/min up to a final bake out at 235 ^{o}C for 5 min. The total run time was about 36 minutes. The mass spectrometer was set to selected ion monitoring (SIM) to detect ions of molecular weight 50 and 52 for methyl chloride, 48 for methyl sulphide, 94 and 96 for methyl bromide, 58 for acetone, 127 and 142 for methyl iodide, 101 and 103 for CFC11, 62 for dimethyl sulphide, 68 and 69 for isoprene and 83 and 85 for chloroform. Peaks corresponding to the above ions on GC-MS traces were identified based on comparison with known retention times for the machine in question (Kelly Redeker, Personal Communication). Peak areas were calculated in arbitrary units

using the GDC Plus Chemistation G1074B software version A.01.00 (© Hewlett-Packard 1989-1996) and then (in the case of pairs of ions for the same volatile) added together. These peak areas were then converted to quantities in parts per trillion using standards and calibrations of trace gases as described previously (Redeker *et al.* 2007; Issitt *et al.* 2022).

2.2.2 TF GC-MS

Pure liquid standards of isoprene (Alfa Aesar) acetone (Sigma-Aldrich), limonene (Sigma-Aldrich) and 2(methylthio)ethanol (Thermo Scientific Chemicals) were initially diluted 1:100 (volume: volume) in LB medium, and 1 mL of this was transferred to a 20 mL glass vial (Gerstel), sealed with a 20 mm Silicone blue/PTFE aluminium crimp cap (Gerstel). For the concentration dropping experiment, an initial dilution of all 4 chemicals at 0.1 M was prepared in LB then serial diluted 6 times in LB. Samples were transferred to 20 mL glass vials as above. Blanks containing 1 mL LB were also included to determine background readings from test samples. Samples were run using the SPME method with a 85 µm thick PAL fiber (Restek) on an Agilent 7890A GC with a Agilent 5975C inert XL MSD with Triple-Axis Detector equipped with either a DB-FATWAX Ultra Inter (Agilent) or Inferno (Zebron) column. For experiments involving dilution series, the LB blanks were run first followed by dilutions in increasing concentration to limit potential carryover effects. For the program that achieved the best separation samples were agitated at 60 °C for 1 min. The run had an initial oven temperature of 40 °C for 2 mins then increased by 10 °C/min to 250 °C, then held for 2 min (total run time 25 min).

2.2.3 GC used for ethylene

2.2.3.1 Calibrations

A gaseous 5000 ppm ethylene/nitrogen mix (BOC) was used to create 6 dilutions (5, 25, 50, 100, 150 and 200 ppm) by volume in 100 mL glass vials (Wheaton). Vials were closed with butyl stoppers and aluminium crimp camps. A neat 5000 ppm ethylene bottle was prepared by sparging a 100 mL glass vial with 5000 ppm ethylene for 2 min then pressurising the bottle with 5000 ppm ethylene to 10 psi. A neat pressurised nitrogen bottle was prepared in the same way. Dilution vials were sparged with nitrogen for 5 min then slightly over pressurised with nitrogen. Vials were then connected to a pressure transducer (Gems™ Sensors & Controllers) via a three-way Luer stopcock connected to a

21G x 1½" Nr 2 Microlance™ (BD) hypodermic needle pierced through the stopper of the bottle (the third outlet was left shut). A second needle was inserted briefly to allow the bottle (which was under positive pressure) to equilibrate to atmospheric pressure then withdrawn. Dilutions were created by withdrawing the desired volume of ethylene and nitrogen from their neat pressurised vials and injecting them into the sparged dilution bottle using airtight SGE glass syringes (10 mL for volumes over 1 mL and 1 mL for volumes under 1 mL). The pressure transducer was used to record pressure changes in the dilution bottles which allowed the exact volume of gas injected to be back calculated. 1 mL samples of each dilution were loaded onto the GC in triplicate as explained in section 2.2.3.6 and integrated to give peak area in AU as explained in section 2.2.3.7. Mean peak area of the triplicate measurements in AU was plotted against ppm of ethylene (adjusted for pressure) to create a calibration curve.

2.2.3.2 Initial ethylene production test

For the initial test plasmids TestC and CtrlC in *E. coli* strain MG1655 were streaked from glycerol stocks on LB agar plates supplemented with chloramphenicol and incubated overnight at 37 ^{o}C. 0.1 mL of chloramphenicol solution (prepared as explained in 2.1.9.14) was added to 7 anaerobic LB broth tubes with resazurin (prepared as explained in 2.1.9.13) and 7 aerobic LB broth tubes (made by adding 5 mL of LB broth to glass test tubes sealed with butyl rubber stoppers and aluminium crimp caps which had been autoclaved at 121 ^{o}C, 15 psi for 20 min). Three individual colonies of each plasmid were inoculated into 3 of each of the tubes (anaerobic tubes were inoculated in the anaerobic chamber). One of each tube was also mock inoculated (i.e. no bacteria transferred). Tubes were incubated overnight at 37 ^{o}C, 150 RPM shaking. On ice, overnight cultures were adjusted to an OD_{600} of approximately 0.05 by first adding 0.1 mL of the overnight culture to a fresh tube of the corresponding type with a 1 mL syringe and 21G x 1½" Nr 2 Microlance™ (BD) hypodermic needle, measuring the OD_{600} on the Spectronic 20D+ spectrophotometer then adjusting by adding more LB medium or overnight culture using syringes and needles if needed. These more concentrated cultures were labelled "A". A second less concentrated "B" version of each culture was made by transferring 0.1 mL of culture A with a syringe and needle into 5 mL of LB medium of the corresponding type. Control cultures were made by transferring 0.1 mL of the overnight controls into corresponding tubes. All cultures were incubated 37 ^{o}C, 150 RPM shaking and growth was

monitored by taking regular OD_{600} measurements. Headspace samples were taken after 12 and 34 hours of incubation by inserting a 21G x 1½" Nr 2 Microlance™ (BD) hypodermic needle attached to a 1 mL gastight glass syringe (SGE) through the butyl stopper of the tube, withdrawing 1 mL headspace, locking the syringe and removing it.

2.2.3.3 Varied antibiotic concentration ethylene test

This was carried out as explained in section 2.2.3.2 with the following amendments. Only TestC in *E. coli* strain MG1655 was used and all LB tubes were aerobic. 10 overnight cultures were prepared by inoculating single *E. coli* colonies into tubes, and an eleventh was prepared by inoculating a smear of several colonies into a tube. Overnight cultures were each subcultured into 4 LB tubes with different final chloramphenicol concentrations (25, 50, 100 and 200 µg/mL) by transferring 0.1 mL of the overnight culture to the recipient tube with needles and syringes, and this was not done on ice.

2.2.3.4 DH5α ethylene test

This was carried out as explained in section 2.2.3.3 with the following amendments. TestC in *E. coli* strain DH5α was used. Four overnight cultures were prepared from individual colonies and one from a smear of multiple colonies. Overnight cultures were all subcultured into fresh LB tubes with 25 µg/mL chloramphenicol.

2.2.3.5 Repeat test of putative ethylene producing cultures

This was carried out as explained in section 2.2.3.4 with the following amendments. Glycerol stocks taken at the end of the initial experiment (section 2.2.3.2) of TestC +O2 A3 and CtrlC -O2 B6 were used to create streak plates. Four overnight cultures from each streak plate were generated by inoculating a smear of multiple colonies in LB tubes.

2.2.3.6 Running headspace samples on GC

1 mL headspace samples held within 1 mL gastight glass syringes (SGE) were loaded into the manual sample port of a Hewlett Packard 5890 Series II GC by inserting a 21G x 1½" Nr 2 Microlance™ (BD) hypodermic needle connected to the syringe through a rubber septum, unlocking the syringe, injecting the contents and pressing "Start" on the control panel. Samples were run on a Restek Rt-Alumina BOND/MAPD PLOT Column, (30 m, 0.32

mm internal diameter, 5 μm thickness) with a nitrogen mobile phase on a 2 min program with the oven temperature set to 105 ^{o}C.

2.2.3.7 Peak integration

Ethylene peaks were integrated within the Clarity™ (version 7.4.0.75) software with all settings on default except integration window set to 1.1 - 1.4 min. For a minority of cases in the traces used to generate the calibration curves, the ethylene peak eluted slightly outside this time window. In these cases, the window was adjusted to a minimum of 1.05 and maximum of 1.7 min. In the small number of cases where more than 1 peak appeared in the integration window, the ethylene peak was identified by comparing its elution time to that of a recent trace in which an ethylene standard had been loaded.

2.3 Sequencing and bioinformatics

2.3.1 16S metagenomic sequencing

16S rRNA Illumina amplicon sequencing was performed to a depth of 50,000 tags per sample with the 341F (CCTAYGGGRBGCASCAG) and 806R (GGACTACNNGGGTATCTAAT) primers for hypervariable region V3 - V4 regions by Novogene, UK. Sequencing was carried out on the Illumina NovaSeq 6000 platform with 250 bp paired end reads. Sequence data processing including paired end read assembly, quality control and chimera removal were performed by Novogene. ASV assignment using DADA2 (Callahan *et al.* 2016), taxonomic assignment using QIIME2 (Bolyen *et al.* 2019) and alpha diversity calculations were all performed by Novogene.

2.3.2 Identification of desirable organisms for isolation from sludge

2.3.2.1 BONCAT metagenomic data

Exact details of the material used to generate the BONCAT data, culture conditions, DNA extraction protocols, sequencing and long read assembly are provided in (Amekan 2021). Briefly: sludge material used to generate the BONCAT data was collected from full scale (≈ 1800 m^{3}) anaerobic digesters at the Yorkshire Water Naburn site and subsequently transferred to a series of laboratory scale 5 L anaerobic digesters at the University of York, which were then starved until activity was undetectable (as determined by volatile fatty acid emissions) prior to addition of treatment nutrients (octanoic acid, synthetic feedstock and water) (Amekan 2021). Samples were taken from triplicate digesters at a series of

time points (0, 6, 12, 18 and 24 hours) post nutrient addition, and active organisms were isolated using a modified BONCAT method. DNA was extracted and sequenced using a combination of long (Oxford Nanopore) and short (Illumina) reads then assembled into contigs (Amekan 2021).

2.3.2.2 Long term digester metagenomic data

Exact details of the material used to generate the standard long term metagenomic dataset, DNA extraction protocols, sequencing and assembly are provided in (Cansdale 2020). Briefly: sludge samples were taken from the same digesters at the Naburn site as used to generate the BONCAT data at approximately two-week intervals over a period of 8 months. DNA was immediately extracted, sequenced using a combination of long (Oxford Nanopore) and short (Illumina) reads, and long reads were assembled into contigs (Cansdale 2020).

2.3.2.3 Assembly of metagenomic data into MAGS and identity assignment

Short reads obtained for the BONCAT (at each of the different timepoints) and long term metagenomic datasets were mapped onto the assembled long reads generated at each corresponding time point and binned into MAGS according to similar changes in contig frequency using an in-house pipeline with the CLUSTard package (github.com/ac1513/CLUSTard) (Cansdale 2020). Organism identity assignments for the MAGS were generated using a Kraken2 (Wood *et al.* 2019) database built using the GTDB taxonomy (Parks *et al.* 2018) (version 9.5).

2.3.2.4 Generating list of "wanted" organisms

An initial list of genera to investigate was produced by using CLUSTard to bin all genera found to be active in all BONCAT conditions. Manual curation was used to remove unsuitable genera (e.g. containing no known genetic tools) leaving 4 candidate genera (*Pseudomonas*, *Bacteroides*, *Bacillus* and *Clostridium*). A list of species to investigate (Table A1, Appendix) was generated by inputting all species belonging to each genus present in the MiDAS database (Nierychlo *et al.* 2020), simple web searches and literature searches (Liu *et al.* 2013; Dong and Zhang 2014; Pyne *et al.* 2014; Lim *et al.* 2017; Martínez-García and de Lorenzo 2017; Charubin *et al.* 2018; Gyulev *et al.* 2018; Joseph *et al.* 2018; García-Bayona and Comstock 2019; McAllister *et al.* 2019). For each species, literature

searches were used to exclude those unlikely to grow in anaerobic conditions. For those that remained, further searches were used to determine whether any genetic tools were available: any organisms for which no means of introducing and expressing exogenous DNA (either through shuttle vectors or chromosomal modification) could be found were not considered further. A subset of 22 organisms from the initial list which met these two criteria were selected as "desirable" organisms (Table A2, Appendix). It is acknowledged that this approach likely generated some false negatives (e.g. organisms incorrectly excluded either because the relevant publication(s) were not found, or this physiology has not yet been adequately described).

2.3.2.5 Generating and querying custom BLAST database: genomes

One genome sequence for each species was first obtained from the NCBI database. If available, a reference or representative genome was obtained (in that order of preference). In cases where these were unavailable, or found to be unsuitable (e.g. split over a large number of contigs), a recent non-representative genome was selected. During this task three species (*Bacteroides vulgatus, Bacillus megaterium* and *Clostridium cellulolyticum)* were found to have been renamed (to *Phocaeicola vulgatus, Priestia megaterium and Ruminiclostridium cellulolyticum,* respectively), although this would not affect downstream analysis. These genome sequences were downloaded from Entrez using the Python package Bio.Entrez (https://biopython.org/docs/1.76/api/Bio.Entrez.html) and concatenated into a single file. This was used to build a BLAST database using the Viking high performance computing servers at the University of York with the command line BLAST+ v2.10.1 (command: makeblastdb -in <sequences.fasta> -dbtype nucl). MAGS were then queried against the custom BLAST database using blastn -task blastn with an evalue cutoff of 0.01 (-evalue 0.01) to produce a tabular output (using outfmt "6 qseqid sseqid stitle pident length mismatch gapopen qstart qend sstart send evalue bitscore").

2.3.2.6 Generating and querying custom BLAST database: 16S

One 16S sequence for each species was obtained from the NCBI database: if available a complete sequence was chosen, if not a recent partial sequence was obtained. All sequences were manually merged into a single fasta file which was used to build a custom BLAST database with Viking using command BLAST+ v2.11.0. MAGS were queried against

this database using blastn -task megablast with an evalue cutoff of 1e-6 (evalue 1e-6, -perc_identity 90) to produce a tablular output (using -outfmt "6 qseqid sseqid pident length mismatch gapopen qstart qend sstart send evalue bitscore stitle").

2.3.3 Pre- and post-THP platewash 16S sequencing and analysis

DNA samples were split into two and pre-amplified by the Technology Facility at the University of York using a nested PCR. The 16S primer pairs 341-b-S-17 (CCTACGGGNGGCWGCAG) / 785-a-A-21 (GACTACHVGGGTATCTAATCC) and Arch-0519-a-S-15 (CAG CMG CCG CGG TAA) / 785-b-A-18 (TACNVGGGTATCTAATCC) were used for the first amplification using NEBNext Q5 Hot start HiFi PCR Master Mix (NEB). The PCR program was 98 °C for 30 sec, 25 cycles of 98 °C for 10 sec, 58 °C for 30 sec and 72 °C for 20 sec then a final step of 72 °C for 2 min. Products were purified with AMPure XP Beads (Beckman Coulter Life Sciences) then a second PCR was carried out with primers 341-b-S-17_ill (TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACCTACGGGNGGCWGCAG) / 785-a-A-21_ill (GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC) using NEBNext Q5 Hot start HiFi PCR Master Mix (NEB) but with a PCR program of 98 °C for 30 sec then 20 cycles of 98 °C for 10 sec, 58 °C for 30 sec, 72 °C for 20 sec then a final step of 72 °C for 2 min. PCR products purified as explained above were sequenced by Novogene using Illumina 250 bp paired end reads with the 341F and 806R primers for hypervariable region V3 - V4. Quality and adapter content were checked using FastQC (v 0.12.1) and MultiQC (v 1.19) (Ewels *et al.* 2016). Sequences were imported to QIIME2 (v 2024.10p) (Bolyen *et al.* 2019), and trimmed using `cutadapt trim-paired` with the adapter sequence AGATCGGAAGAG. DADA2 was used to denoise paired end sequences with parameters --p-trunc-len-f 0 --p-trunc-len-r 0. A phylogenetic tree was generated with parameter `phylogeny align-to-tree-mafft-fasttree`. A rarefaction curve was generated with `--p-max-depth`, median frequency per sample set to (A:91454, B:86503). Taxonomic classification was performed using QIIME2 (Bolyen *et al.* 2019) with parameters set to `feature-classifier classify-sklearn` using the GreenGenes (DeSantis *et al.* 2006) 2024.09 full length database. Taxa was collapsed to genera level using `taxa collapse` and `--p-level 6`. The result was converted to a BIOM then a TSV file for analysis.

2.3.4 Eight organism test pool sequencing, assembly and analysis

Pooled and separate gDNA samples derived from organisms A - H were submitted to Novogene for shotgun metagenomic sequencing using the Illumina Novaseq X Plus platform with 150 bp paired end reads with 5 Gb of raw data per sample. Unassembled sequence data was taxonomically assigned using Kraken2 with the GTDB_r89_54k database. Data was visualised and Sankey plots generated using the browser version of the Pavian metagenomics data explorer (Breitwieser and Salzberg 2020). The short reads were trimmed using Cutadapt (version 4.2) (Martin 2011), assembled using metaSPAdes (Nurk *et al.* 2017), and quality of assemblies was inspected using SeqKit (Shen *et al.* 2016). Assemblies were converted to BAM files, binned using MetaBAT (Kang *et al.* 2019) and MaxBin (Wu *et al.* 2016) and the outputs of both rebinned using DAStool (Sieber *et al.* 2018). Taxonomic assignment of bins was performed using Kraken2 with the GTDB_r89_54k database. As Kraken2 assumes a uniform read length, the MAGS were then run against Kraken2 outputs using a custom python script to normalise the length so that the proportion of the MAG bases assigned to each annotation were accounted for.

2.3.5 96 well plate sequencing and analysis

The 20 pooled row and column gDNA samples were long read sequenced by the Technology Facility at the University of York. Samples were barcoded with the Oxford Nanopore Native Barcoding Kit (SQK-NBD114-24) and sequenced using a promethION 24 running MinKNOW version 24.11.11 with base calling through Dorado version 7.6.8 (Oxford Nanopore Technologies). Raw sequence data was first concatenated into separate files corresponding to each of the 20 original samples using the barcodes. The quality of the raw data was checked using FastQC and MultiQC. The data from each row and column was then assembled with Flye (Kolmogorov *et al.* 2019) (version 2.9.6) using –nano-hq. The row files were merged and used to make a custom BLAST database of the row assemblies. Column assemblies were queried against this database, the outputs were sorted by bitvalue, evalue and percentage match and the best match for each column and row assembly was assigned to the corresponding intersection in the array. A custom python code was used to generate a fasta file containing all the contigs for each of the 96 intersections in the array. Taxonomic assignment was performed on the intersections with the NCBI nucleotide BLAST database (dated 18th Dec 2024) and using Kraken2 with the r89 GTDB 64gb database. Taxonomic assignments for the intersections are based on

Kraken2 identity and were only included for contigs that matched the r89 GTDB 64gb database for >10% of their length.

2.3.6 Selective plating isolate whole genome sequencing and identification

Separate gDNA samples from each isolate were submitted to BMKgene for shotgun metagenomic sequencing using the Illumina NovaSeq X Plus platform with 150 bp paired end reads with 5 Gb of raw data per sample. Reads were taxonomically assigned using Kraken2 with the GTDB_r89_54k database. Data was visualised and Sankey plots generated using Pavian.

2.3.7 Experimental evolution sequencing and SNP calling

gDNA samples from the 8 directed evolution *E. coli* strain MG1655_rLP5(nth-ydgR) populations, the ancestral population control and strain MG1655 without the RFP insertion were submitted to Novogene for shotgun metagenomic sequencing using the Illumina Novaseq X Plus platform with 150 bp paired end reads with 5 Gb of data per sample. Cutadapt (version 5.1) was used to remove sequencing adapters and filter reads to a q score of 20. DNA sequences for each population were aligned against the NCBI *E. coli* K-12 MG1655 reference genome GCF_000005845.2 using BWA version 0.7.17-r1188 (Li and Durbin 2009). Picard (version 2.25.5) was used to add read groups and mark duplicates. Single nucleotide polymorphisms (SNPs) were identified in the populations vs the reference genome using FreeBayes version 1.3.6 (Garrison and Marth 2012). The VCF file output was filtered using VCF tools (version v0.1.16) (Danecek *et al.* 2011) with parameters: --mac 1, --minQ 30, --max-missing 1 to display the SNPs of interest. These were manually curated to remove any SNPs that were shared by all ancestral and experimental samples, or were only present in the reference genome. The remaining SNPs were run using SNPeff (version 5.4a) (Cingolani *et al.* 2012) against the SNPeff database for Escherichia_coli_str_k_12_substr_mg1655_gca_000005845. SNPs with extreme strand bias (SAF/SAR metrics) or read placement bias (RPR/RPL), and therefore likely erroneously identified due to sequencing/alignment artefacts, were manually removed. Allele identities and corresponding values were inverted in cases where the ancestral strain was annotated as “ALT” such that the ancestral status and values for all alleles were set to correspond to the ancestral population rather than the reference genome. Allele frequencies for each population were calculated from values in the VCF file by dividing

the alternative allele observation count (AO) by the sum of AO and the reference allele observation count (RO) then multiplying by 100.

2.4 Miscellaneous

2.4.1 Figures

Figures were created in Graphpad, Microsoft Excel, R (R Core Team 2025) using RSudio (Posit team 2025) or BioRender (https://BioRender.com).

2.4.2 Statistics

The following packages were used in R for statistics and figure production: Tidyverse (Wickham *et al.* 2019), ggplot2 (Wickham 2016), dplyr (Wickham *et al.* 2023), emmeans (Lenth and Piaskowski 2025). Post-hoc testing for ANOVA was performed using Tukey's Honestly Significant Difference test (Tukey 1949).

Chapter 3: Whole Cell Volatile Bioreporters

3.1 Introduction

3.1.1 Context of work

The overall aim of this PhD project is to investigate biological approaches for monitoring and manipulating the mixed microbial communities which underpin AD. In this chapter an attempt was made to develop a tool for monitoring in AD communities: a whole cell volatile bioreporter.

3.1.2 Repurposing an existing system

The approach selected was to adapt an existing quantitative volatile reporter system (see section 1.20, General Introduction) for use in AD. In the chosen system, bacterial cells had been engineered to produce the volatile methyl bromide in response to an inducer, and constitutively produce ethylene gas to enable signal normalisation (Cheng *et al.* 2018). The genetic components required for this system were expressed chromosomally in *E. coli* strain MG1655, and from a plasmid (pHC01) in the bacterium *Shewanella oneidensis* (Cheng *et al.* 2018). While the methyl halide transferase gene used to produce methyl bromide in the above system has been demonstrated to work anaerobically, the ethylene producing gene used requires oxygen (Cheng *et al.* 2016). Since a minimum of two gases are required to enable normalisation to biosensor cell number (Cheng *et al.* 2018), a minimum of one additional appropriate volatile must be identified to enable this system to be adapted for use in an anaerobic environment.

3.1.3 Identifying additional candidate volatiles

Volatiles suitable for use in this reporter system are subject to several requirements. Firstly, they must not already be present in the headspace of a typical anaerobic digester in significant quantities since this would make the reporter generated signal indistinguishable from background. Secondly, genes capable of producing enzymes which generate these volatiles must exist in nature, and must be able to function in the chosen chassis organism. The volatiles produced must have between one and six carbon atoms due to the specification of the GC-MS equipment available for this project (Kelly Redeker, personal communication). Finally, the genetic pathways required to produce these

volatiles should be as short as possible to limit energetic demands on the host, as well as necessitating the exogenous expression of fewer genes. The GC-MS available for this project was used routinely to quantify several volatiles, so these may represent a sensible starting point for identifying additional anaerobically producible candidate volatiles.

3.1.4 Anaerobic ethylene synthesis

Ethylene is one of the two volatiles which had already been used in the dual gas reporter system, however the pathway used to produce it was reliant on oxygen (Cheng *et al.* 2018), (bottom left, Figure 3.1). Since this work shows ethylene is suitable as a reporter, one possibility would be to attempt to generate this volatile anaerobically instead of using a different second gas. The majority of routes for biological ethylene synthesis are involved with methionine salvage and many require oxygen (North *et al.* 2017) (central circle, Figure 3.1). To the author's knowledge there are only two anaerobic biological pathways for ethylene synthesis. One involves the promiscuous activity of nitrogenases (Benton *et al.* 2001), but requires exogenous addition of acetylene (Bytnerowicz *et al.* 2019) (bottom right, Figure 3.1), which would make this unsuitable for use in a bioreporter. The other is a recently reported branch of the methionine salvage pathway found in some phototrophic bacteria (North *et al.* 2017; North *et al.* 2020a) (red arrows, Figure 3.1) which has not yet been expressed exogenously. Commensal strains of *E. coli* such as MG1655 possess the metabolic capacity to produce up to the compound MTR (green arrows, Figure 3.1). Some extraintestinal pathogenic strains (ExPEC) such as ATCC 25922 can progress further up to 2(methylthio)ethanol (North *et al.* 2020b) (blue arrows, Figure 3.1), while some phototrophic bacteria such as *Rhodospirillum rubrum (R. rubrum)* can additionally progress the remainder of the way back to methionine (North *et al.* 2020a; North *et al.* 2020b) (red arrows, Figure 3.1). This presents the intriguing possibility that by heterologous expression of different pathway components from plasmids in different *E. coli* strains, it may be possible to produce both ethylene and the volatile intermediate 2(methylthio)ethanol, which could be useful for gas reporting. It is noteworthy that anaerobic biosynthesis of ethylene would potentially also be useful it its own right since it is a valuable chemical precursor (see section 1.23.4, General Introduction).

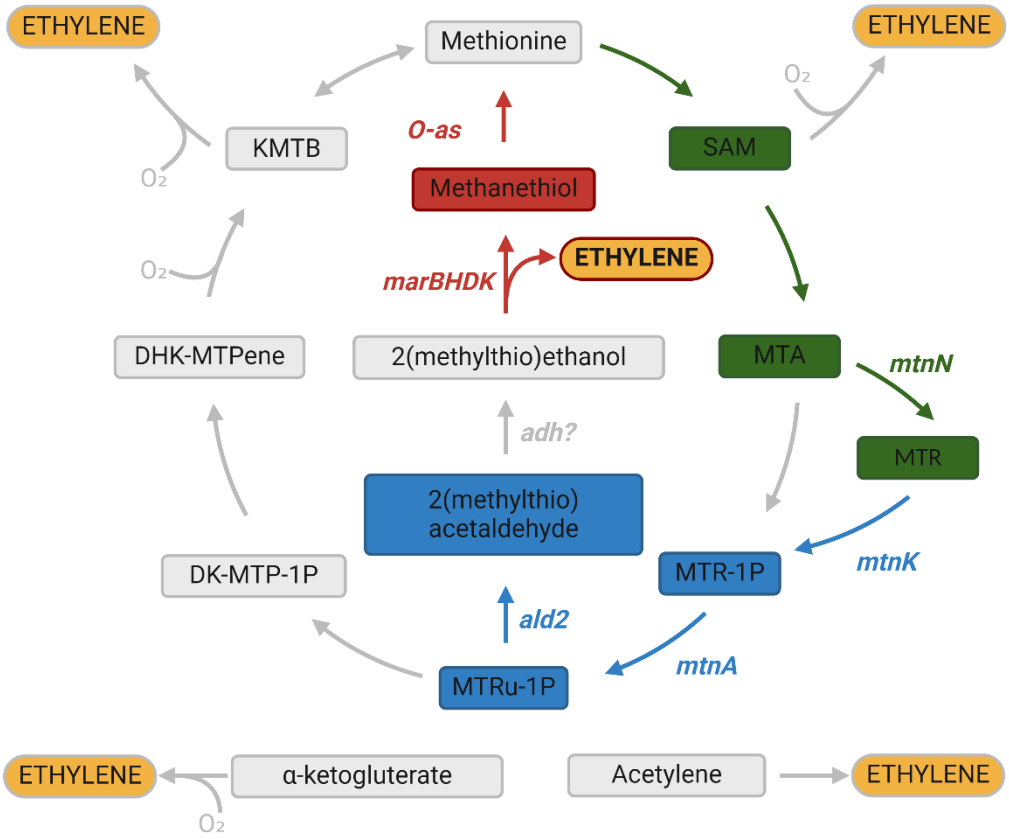


Figure 3.1: Ethylene synthesis pathways.
The majority of modes of ethylene synthesis are related to the methionine salvage pathway (central circle) and require oxygen. Ethylene can also be synthesised by a dedicated ethylene forming enzyme (bottom left) and through the promiscuous activity of nitrogenases (bottom right). A strictly anaerobic bacterial ethylene synthesis pathway has recently been discovered (coloured boxed and arrows). Boxes: compounds. Italics: enzymes. Adapted from (North *et al.* 2017; North *et al.* 2020a).

3.1.5 Aims

The aim of this chapter is to repurpose the previously described quantitative whole cell volatile bioreporter described in (Cheng *et al.* 2018) for use in anaerobic environments such as AD. This can be split into the sub aims:

1. Select additional candidate volatiles which can be biosynthesised anaerobically.
2. Obtain plasmid constructs suitable for producing selected volatiles anaerobically.
3. Validate biosynthesis of chosen volatiles in *E. coli.*
4. Modify the existing system by replacing ethylene with the best anaerobically produced volatile candidate and test the system in AD.

3.2 Results

3.2.1 Select additional candidate volatiles which can be biosynthesised anaerobically

3.2.1.1 Rational exclusion of volatiles already detectable on initially available GC-MS.

A literature search was first used to assess the suitability for anaerobic gas reporting of the nine volatiles which the GC-MS available in this project was already calibrated and routinely used to detect. The primary criteria used to assess suitability was whether the gene(s) for the enzyme(s) which synthesise each volatile (or those responsible for synthesis of nearby chemical precursors) should theoretically work anaerobically (e.g. not directly requiring oxygen), and whether these enzyme(s) have already been expressed exogenously. Six of the nine candidates did not meet these criteria and were excluded (Table 3.1).

Table 3.1: Suitability of existing GC-MS volatiles for anaerobic gas reporting.
Result of literature investigation into the suitability of volatiles that the available GC-MS was already calibrated to quantify for anaerobic gas reporting. Where a volatile was deemed to be unsuitable, the reasoning is given in the notes.

Volatile	Should Work Without Oxygen?	Previously Engineered?	Suitable?	Notes
Methyl bromide	✓ (Cheng *et al.* 2016)	✓ (Cheng *et al.* 2016)	✓	
Methyl chloride	✓ (Cheng *et al.* 2016)	✓ (Cheng *et al.* 2016)	X	Uses same enzyme as methyl bromide (Cheng *et al.* 2016).
Methyl iodide	✓ (Cheng *et al.* 2016)	✓ (Cheng *et al.* 2016)	X	Uses same enzyme as methyl bromide (Cheng *et al.* 2016).
Methane thiol	✓ (Schäfer and Eyice 2019)	X	X	Not previously engineered.
Dimethyl sulphide	✓ (Schäfer and Eyice 2019)	X	X	Not previously engineered.
Isoprene	✓ (Aldridge *et al.* 2021)	✓ (M. Li *et al.* 2018)	✓	
Acetone	✓ (Frey *et al.* 2018)	✓ (Bermejo *et al.* 1998)	✓	
Chloroform	?	X	X	Synthesised by some organisms (Laturnus *et al.* 2002), enzymes not known.
CFC11	X	N/A	X	Synthetic chemical: no biological source (McFarland and Kaye 1992).

3.2.1.2 Additional volatiles not already detectable on available GC-MS

Three additional volatiles (ethylene, 2(methylthio)ethanol and limonene) that the GC-MS was not calibrated to detect, but which should be detectable on it according to their chemical properties and which met the inclusion criteria that had been used to select acetone, isoprene and methyl bromide were also added to the list of candidate anaerobic reporting gases.

3.2.2 Obtain plasmid constructs suitable for producing selected volatiles anaerobically

3.2.2.1 Ethylene and 2(methylthio)ethanol

3.2.2.1.1 Pathway engineering rationale

By heterologously expressing different pathway components from a recently described anaerobic ethylene synthesis pathway on plasmids in different *E. coli* strains (Figure 3.1) it may be possible to produce both ethylene and the volatile intermediate 2(methylthio)ethanol, which could be useful for gas reporting (Table 3.2).

Table 3.2: Proposed *E. coli* strain/gene combinations for ethylene synthesis.

Construct	*E. coli* Strain	Genes to Introduce	Expected Volatile
TestA	MG1655	*mtnK*, *mtnA*, *ald2*	2(methylthio)ethanol
TestB	ATCC 25922	*marB*DHK, *O-as*	Ethylene
TestC	MG1655	*mtnK*, *mtnA*, *ald2*, *marBDHK*, *O-as*	Ethylene

3.2.2.1.2 Obtaining genes

Appropriate gene sequences for the pathway components were first obtained. The *mtnA*, *mtnK* and *ald2* gene sequences from *E. coli* ATCC 25922 were obtained from NCBI. The sequences for *marB*, *marD*, *marH* and *marK* from *R. rubrum* were obtained from (North *et al.* 2020a).

The *O-as* gene was less straightforward as it was found to have four homologues in *R. rubrum* according to the KEGG database (Kanehisa *et al.* 2023). It is not known whether all four genes are essential, or whether they are redundant and a single gene would be sufficient to perform the enzyme function. It was preferable to avoid introducing more than one additional gene on top of the existing seven gene pathway already required unless absolutely necessary to limit excessive metabolic burden on host cells.

To determine which of the *R. rubrum O-as* genes to include in the constructs a series of protein alignments were performed to investigate sequence similarity between orthologous genes in organisms either experimentally validated to possess a functional version of this anaerobic ethylene synthesis pathway (*R. palustris* strain CGA010 and *B.*

viridis), or experimentally validated to not (*R. capsulatus*) (Figure 3.2). These three organisms were found to have seven, three and one version of this gene, respectively. The single orthologue from *R. capsulatus*, which is not capable of this pathway, was found to cluster with Rru_A2075, suggesting that this gene is not required. This was unsurprising since protein alignments of the four *R. rubrum* genes had suggested that this gene was substantially larger than the others (612 amino acids compared with a mean of 437 for the other three). Rru_0774 clustered with one gene from *R. palustris* but not from *B. viridis*, suggesting this gene may not be essential. Genes Rru_0784 and Rru_1836 clustered with a single gene from both *R. palustris* and *B. viridis*. This pair of genes had also been found to be the most similar to each other at the protein level of the four *R. rubrum O-as* genes. Rru_0784 (hereafter simply O-as) was selected because it had been found to be one of the most highly upregulated genes under sulphur starvation conditions in *R. rubrum* (i.e. the condition this pathway is biologically employed) (North *et al.* 2020a).

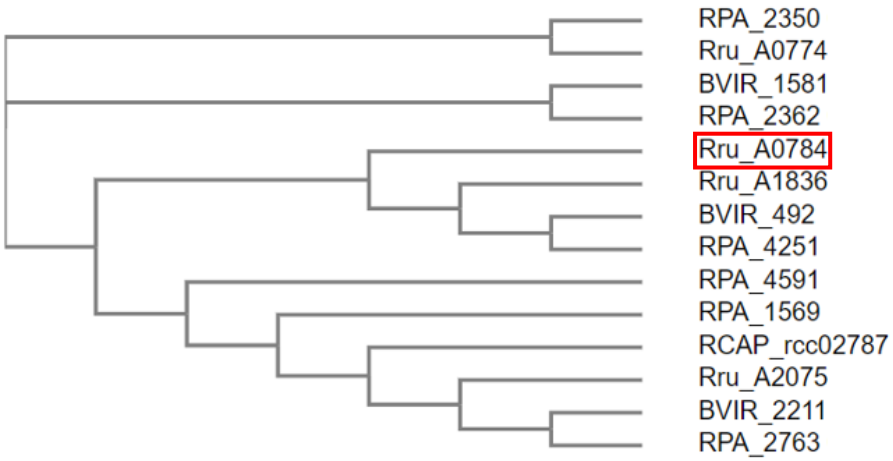


Figure 3.2: O-acetylhomoserine sulfhydrylase (O-as) alignments.
Protein sequence similarity tree of the O-acetylhomoserine sulfhydrylase genes from *R. rubrum* (Rru), *R. palustris* (RPA), *B. viridis* (BVIR) and *R. capsulatus* (RCAP). The selected gene is shown in the red box.

3.2.2.1.3 Assembly

Since the desired constructs contained several genes and were combinatorial in nature, a Golden Gate assembly system based on the Start-Stop kit (Taylor *et al.* 2018) was used. All eight selected gene sequences were codon optimised for expression in *E. coli* and adapted for use with the kit by removing restriction sites that would interfere with assembly (Taylor *et al.* 2018). An assembly scheme (Figure A1, Appendix) was designed and tested *in silico* to produce the series of plasmids described in Table 3.2. The same

weak constitutive promoter (J23116) and RBS (RBSc42), and a strong terminator (ECK120019600) provided with the Stop-Start kit were used for all genes.

The plasmid TestA contains the *mtnK*, *mtnA* and *ald2* genes (Figure 3.3) which, when introduced into a commensal *E. coli* strain such as MG1655 should enable it to convert the dead-end metabolite MTR to 2(methylthio)acetaldehyde, which should then be converted to 2(methylthio)ethanol from endogenous alcohol dehydrogenase(s). The control plasmid CtrlA lacks *mtnK* so should be incapable of progressing past MTR. The inclusion of the other two genes ensures this plasmid should impart a similar metabolic burden as TestA onto host cells, serving as a more comparable control.

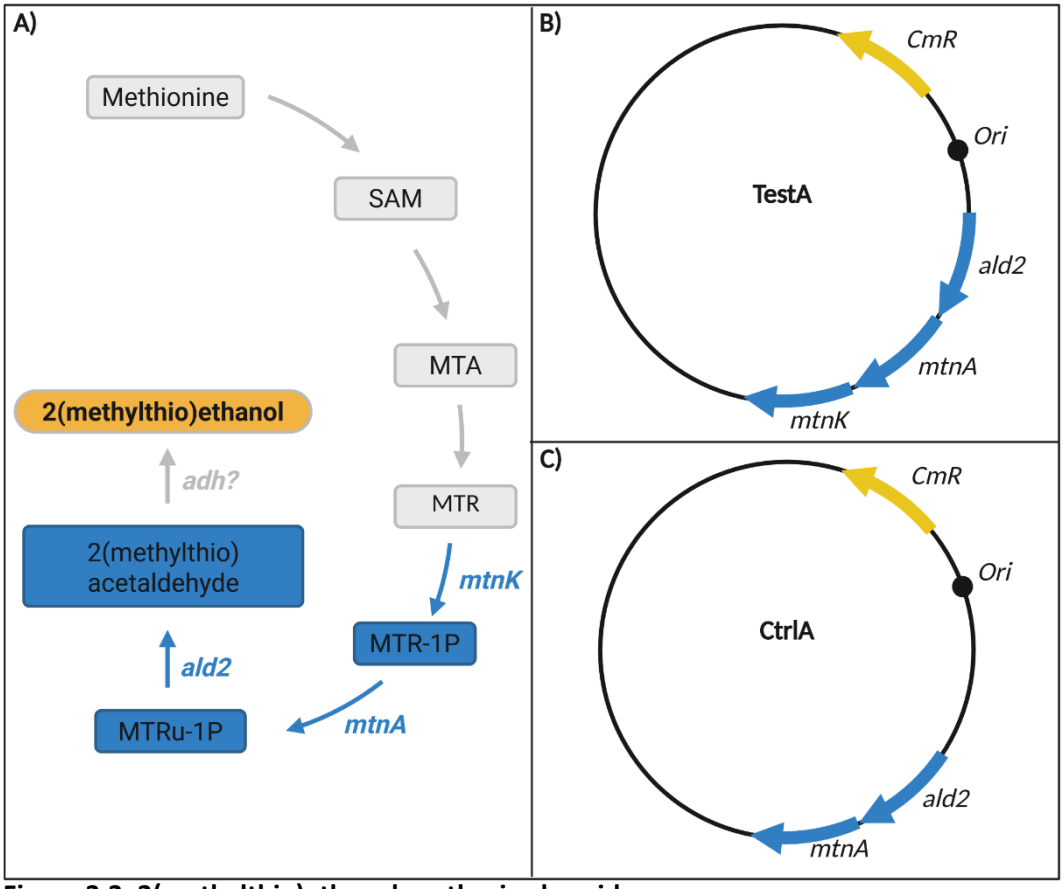


Figure 3.3: **2(methylthio)ethanol synthesis plasmids.**
A) The 2(methylthio)ethanol precursor MTR is synthesised endogenously in commensal *E. coli*. Introduction of the genes *mtnK*, *mtnA* and *ald2* theoretically enables commensal *E. coli* to synthesise 2(methylthio)ethanol. Boxes: compounds. Italics: enzymes. Endogenous enzymes and compounds are grey, exogenously introduced ones are coloured. **B)** TestA should enable *E. coli* to produce 2(methylthio)ethanol. **C)** CtrlA lacks *mtnK* so should not enable *E. coli* to produce 2(methylthio)ethanol. For clarity constitutive promoters, ribosome binding sites and terminators are now shown. Ori = p15A *E. coli* replication origin. *CmR* = chloramphenicol resistance gene.

Plasmid TestB contains the four gene *marBHDK* cluster which was intended to impart the ability for *E. coli* strain ATCC 25922, which is already capable of producing 2(methylthio)ethanol, to produce ethylene (Figure 3.4). The inclusion of the *O-as* gene completes the pathway, which should theoretically increase ethylene yields by enabling it to function as a cycle (i.e. instead of methanethiol being a dead-end product) (Figure 3.4). The control plasmid CtrlB was designed using the sample principle as described for CtrlA.

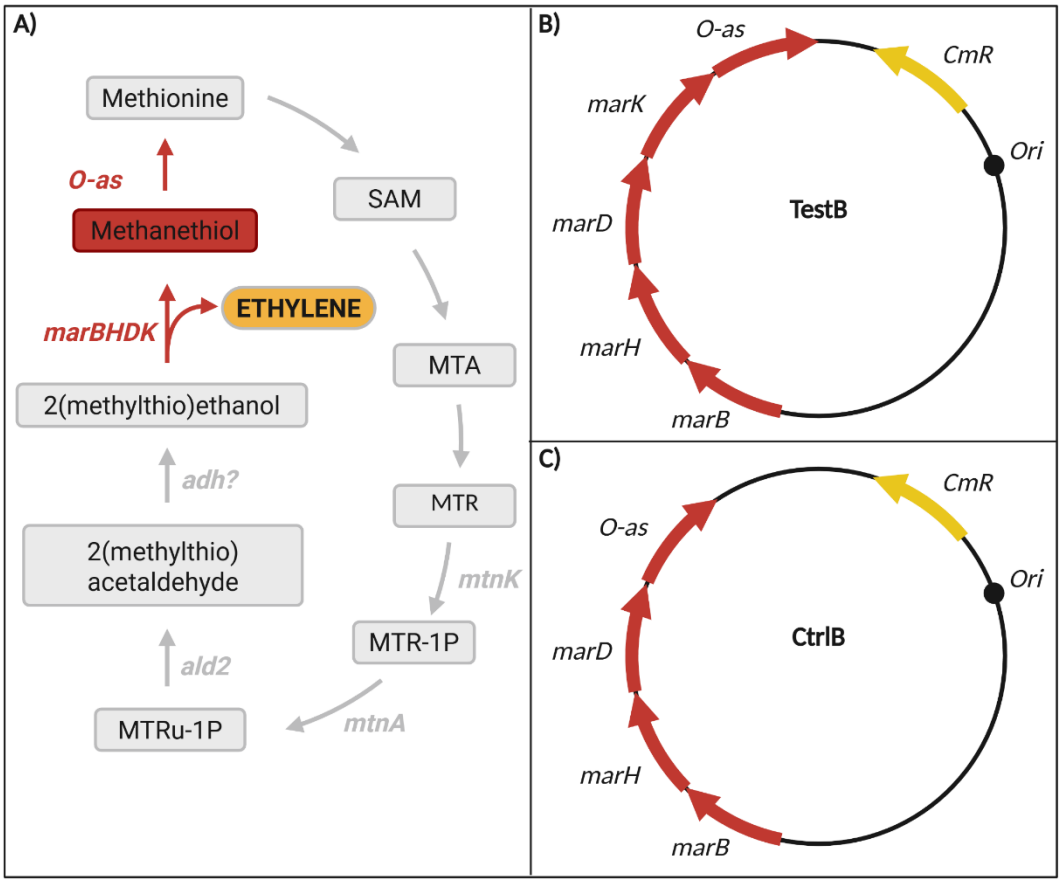


Figure 3.4: Ethylene synthesis plasmids (partial pathway).
A) The ethylene precursor 2(methylthio)ethanol is synthesised endogenously in *E. coli* strain ATCC 25922. Introduction of the genes *marBHDK* and *O-as* theoretically enables *E. coli* ATCC 25922 to synthesise ethylene. Boxes: compounds. Italics: enzymes. Endogenous enzymes and compounds are grey, exogenously introduced ones are coloured. **B)** TestB should enable *E. coli* ATCC 25922 to produce ethylene. **C)** CtrlB lacks *marK* so should not enable *E. coli* ATCC 25922 to produce ethylene. For clarity constitutive promoters, ribosome binding sites and terminators are now shown. Ori = p15A *E. coli* replication origin. *CmR* = chloramphenicol resistance gene.

Construct TestC contains all eight genes from the anaerobic ethylene synthesis pathway (Figure 3.5) and should enable commensal *E. coli* to synthesise ethylene. A control plasmid (CtrlC) lacking one essential gene was also designed (Figure 3.5).

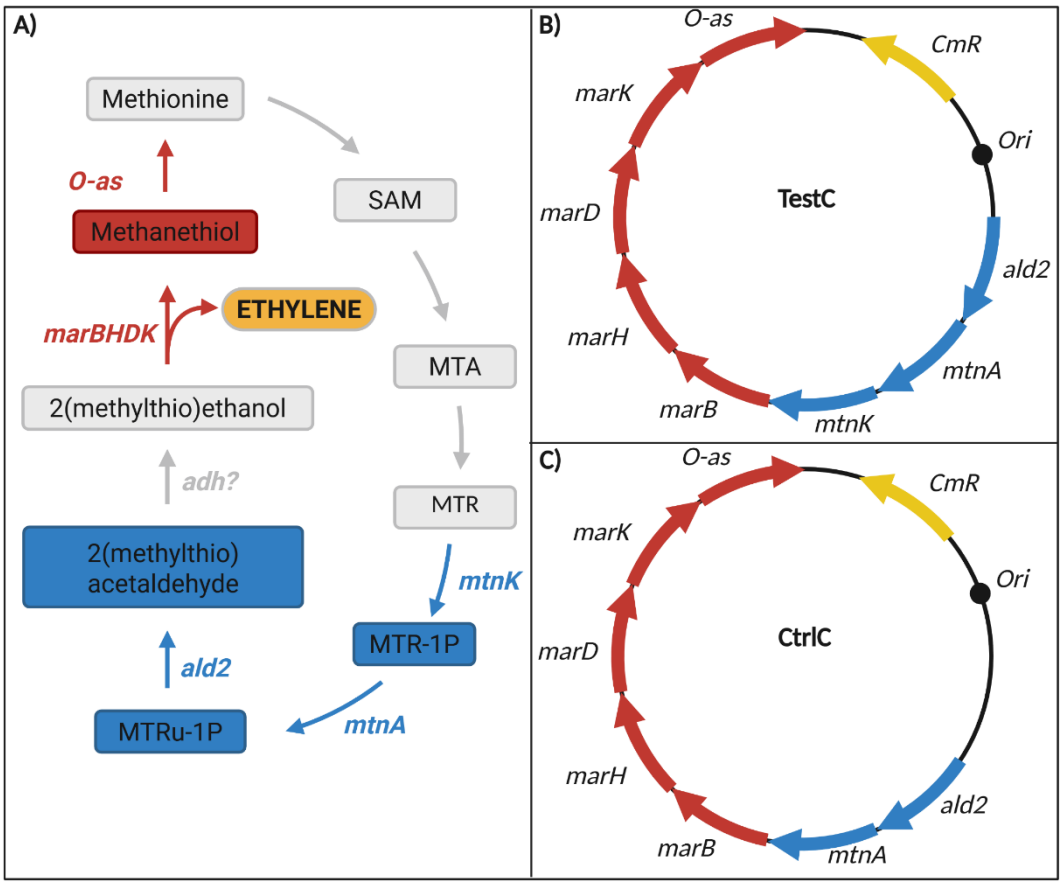


Figure 3.5: Ethylene synthesis plasmids (full pathway).
A) The ethylene precursor MTR is synthesised endogenously in commensal *E. coli*. Introduction of the genes *mtnK*, *mtnA*, *ald2*, *marBHDK* and *O-as* theoretically enables commensal *E. coli* to synthesise ethylene. Boxes: compounds. Italics: enzymes. Endogenous enzymes and compounds are grey, exogenously introduced ones are coloured. **B)** TestC should enable *E. coli* to produce ethylene. **C)** CtrlC lacks *mtnK* so should not enable *E. coli* to produce ethylene. For clarity constitutive promoters, ribosome binding sites and terminators are now shown. Ori = p15A *E. coli* replication origin. *CmR* = chloramphenicol resistance gene.

The eight genes were ordered as synthetic DNA fragments and sequentially assembled according to the designed scheme (Figure A1, Appendix) to produce the six anaerobic ethylene pathway plasmids described above. At each stage of assembly, plasmid intermediates were verified by a combination of colony PCR, diagnostic digestion and full plasmid sequencing. Only plasmids in which all functional components were found to have the expected sequences were taken forward to the next stage of assembly. All six

plasmids were successfully assembled, and each was confirmed to have the expected DNA sequence according to earlier *in silico* predictions.

3.2.2.2 Acetone

Acetone is anaerobically synthesized naturally by solventogenic *Clostridia* (Frey *et al.* 2018) using a three enzyme pathway containing a thiolase (thl), coenzyme A transferase (ctfAB) and acetoacetate decarboxylase (adc) (Nielsen *et al.* 2009) (Panel A, Figure 3.6). This pathway has previously been exogenously expressed in *E. coli* (Bermejo *et al.* 1998). A plasmid containing these genes arranged as an operon under the control of a constitutive promoter was found on the Addgene repository. Plasmid pETM6_Ptrc_BsMdh_BmHps_BmPhi_Pthl_CaThl_CaCtfAB_CaAdc, hereafter pETM6_Ptrc, contains several additional genes not required for this purpose (Panel B, Figure 3.6) (Bennett *et al.* 2018). A PCR scheme for generating a control plasmid (Panel C, Figure 3.6) which has the same composition as the original except for the absence of the thiolase gene was designed.

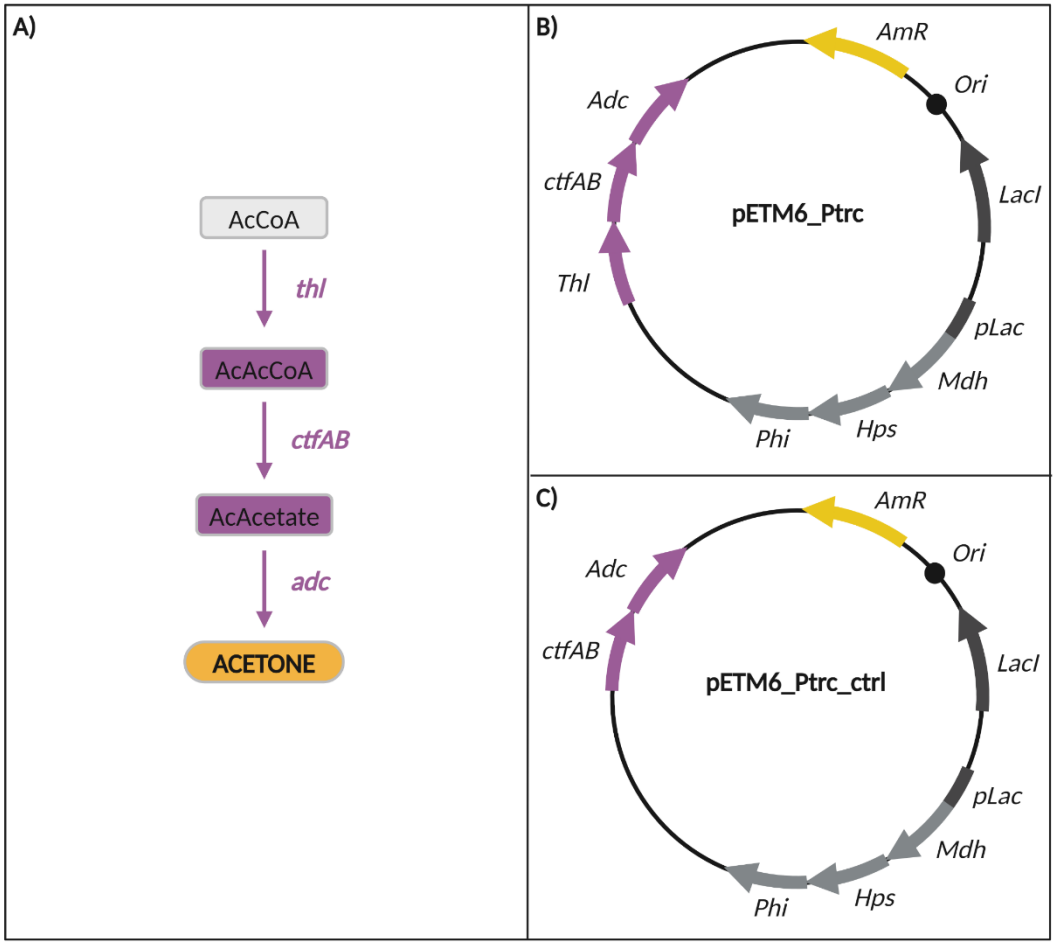


Figure 3.6: Acetone synthesis plasmids.
A) Introduction of the genes *thl*, *ctfAB* and *adc* should enable *E. coli* to synthesise acetone from the glycolysis product acetyl-coA. Boxes: compounds. Italics: enzymes. Endogenous enzymes and compounds are grey, exogenously introduced ones are coloured. Adapted from (Nielsen *et al.* 2009). **B)** pETM6_Ptrc should enable expression of acetone in *E. coli* due to the inclusion of the *adc*, *ctfAB* and *thl* genes. The *Mdh*, *Hps* and *Phi* genes are not involved in acetone synthesis. **C)** pETM6_Ptrc_ctrl lacks *thl* so should not enable *E. coli* to produce acetone. For clarity constitutive promoters, ribosome binding sites and terminators are not shown. Ori = Coli E1 *E. coli* replication origin. *AmR* = ampicillin resistance gene.

pETM6_Ptrc pDNA was extracted from the *E. coli* strain it was delivered in and the plasmid was sequenced and found to be identical to that provided by Addgene. The plasmid was transformed into *E. coli* strain MG1655 for expression. Diagnostic digestions of the pDNA prior to and after transfer to MG1655 had identical banding patterns (data not shown), suggesting that the intact plasmid is present in MG1655 and is ready for testing.

3.2.2.3 Other volatiles

Plasmids suitable for the synthesis of limonene (Alonso-Gutierrez *et al.* 2013) and methyl bromide (Fulk *et al.* 2022) were obtained and transformed into *E. coli* MG1655 for expression, with control plasmids each lacking one gene from each pathway designed as described for acetone (Figure 3.6). A plasmid suitable for expressing isoprene was obtained by ordering a codon optimised version of the isoprene synthase gene (IspS) (M. Li *et al.* 2018) and assembling it using the same kit and a similar process as explained for the ethylene plasmids (Figure A1, Appendix).

3.2.3 Validate biosynthesis of chosen volatiles in *E. coli*

3.2.3.1 Acetone

The strain validation was started with acetone as this gas was one of the most routinely detected on the GC-MS available, and the plasmid obtained for its expression (pETM6_Ptrc) did not require induction and had already been validated to work in *E. coli*. Three separate four point calibration curves were generated for acetone detection by using the GC-MS (Panels A and B, Figure A2, Appendix), however when plotted together these gave an unacceptably low R^2 value (Panel C, Figure A2, Appendix).

While calibration data is required to translate readings from the GC-MS from arbitrary units (AU) to concentrations, it is not required to estimate relative levels of volatiles detected (as relative levels can be calculated simply by dividing the reading the sample gives in AU for a given volatile by the volume of gas that sample was loaded in). Given the difficulties in generating reproducible calibrations for acetone, this relative method was used to compare the level of acetone generated from headspace samples of *E. coli* MG1655 containing the acetone producing pETM6_Ptrc plasmid vs control cultures of MG1655 without plasmid (Figure 3.7).

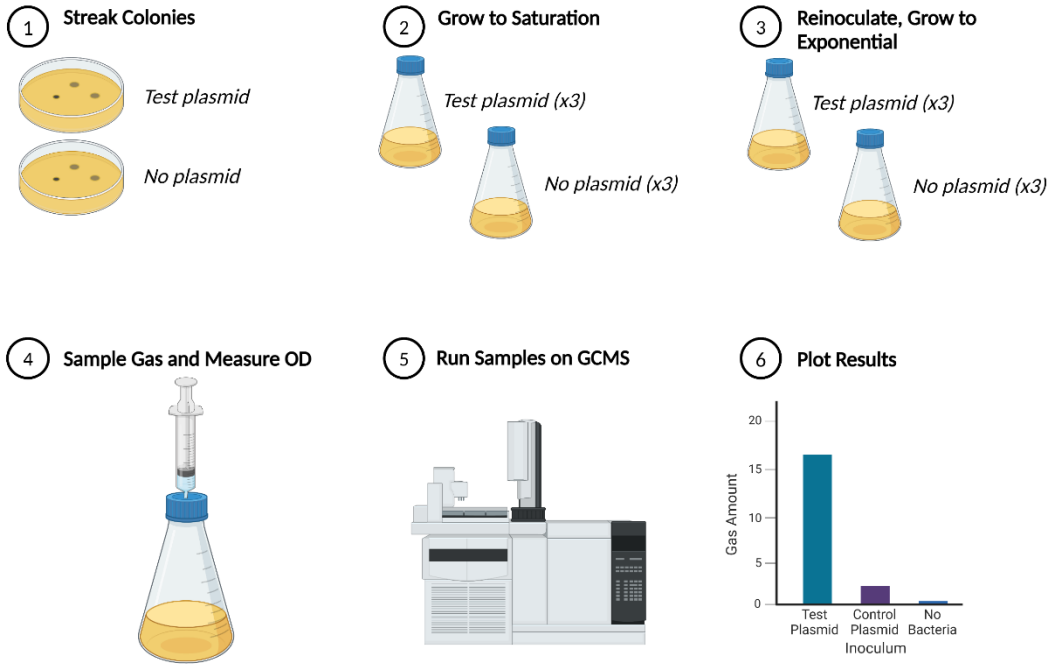


Figure 3.7: Overview of acetone plasmid validation experiment.
1) *E. coli* MG1655 carrying pETM6_Ptrc (Test plasmid) which should be able to produce acetone, and a control containing no plasmid are streaked on agar plates. **2)** Colonies are picked in triplicate and grown in broth to saturation to generate starter cultures. **3)** Starter cultures are reinoculated into broth in airtight vessels and grown to exponential phase. **4)** Headspace gas samples are taken from each vessel at exponential growth phase. **5 – 6)** Samples are run on the GC-MS to determine relative levels of acetone and results are plotted.

As expected, reinoculated cultures (Stage 3, Figure 3.7) carrying the acetone producing plasmid grew slightly slower than those without (Figure 3.8), presumably due to the metabolic burden of the plasmid on the cells. Despite this, both cultures reached a similar OD_{600} at 10 hours post inoculation at which point headspace gas was sampled for acetone analysis on the GC-MS.

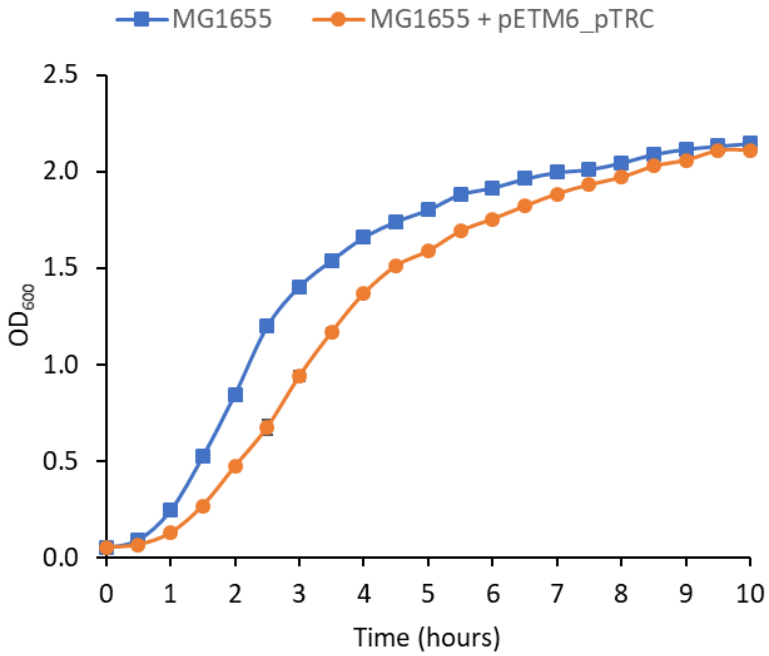


Figure 3.8: Growth dynamics of *E. coli* carrying pETM6_Ptrc. Growth curves for *E. coli* carrying plasmid pETM6_Ptrc for acetone production (orange points) or control *E. coli* carrying no plasmid (blue points). Point are means of triplicate cultures. Error bars are +/- one standard deviation.

On analysis, issues with the machine resulted in the loss of one of the triplicate control samples (MG1655, Figure 3.9). While there was a marginal (approximately 2-fold) higher level of acetone detected in the plasmid bearing samples, this difference was not statistically significant (Figure 3.9).

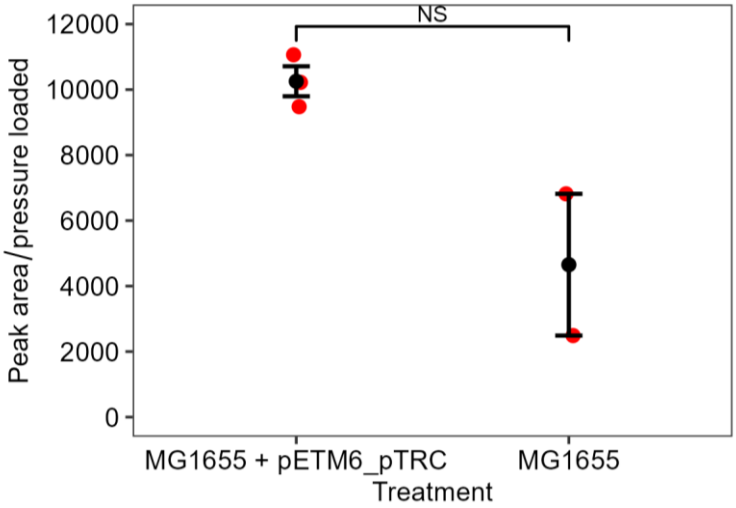


Figure 3.9: Relative acetone readings for *E. coli*.
Mean relative acetone readings were not significantly different between headspace samples from *E. coli* strain MG1655 carrying plasmid petM6_pTRC and control *E. coli* without plasmid (Welch two sample t-test: t = 2.5323; d.f. = 1; p = 0.2224). Error bars are +/- one standard error.

To investigate whether the bacteria and/or plasmid not working as intended may have been responsible for this unexpected result, or whether it may have been due to an issue detecting acetone on the GC-MS, a second similar experiment was carried out, but this time using a known quantity of acetone spiked into the headspace of culture vessels vs control vessels containing lab air (Figure 3.10).

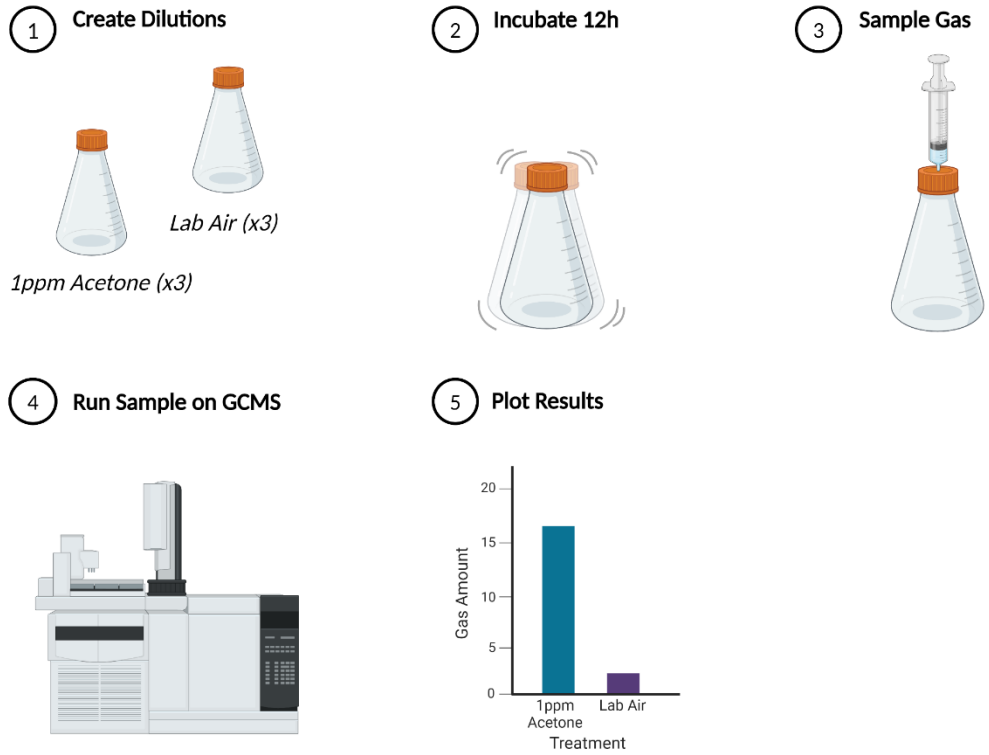


Figure 3.10: Overview of acetone spiking experiment.
1) Triplicate 1 ppm acetone dilutions are made in airtight vessels. Control vessels without acetone are also prepared. **2)** Vessels are incubated for 12 hours with shaking to mimic culture growth. **3)** Headspace gas samples are taken from each vessel. **4-5)** Samples are run on the GC-MS to determine relative levels of acetone and results are plotted.

Higher end estimates for the concentration of acetone in ambient air are around 300 times lower than the 1 ppm included in the spiked bottles (Singh *et al.* 1994), so a substantial difference in relative acetone readings was expected between the two treatments. While a marginally lower mean relative level of acetone was found in the headspace of control samples (921 +/- 474 standard error (SE)) compared with the spiked samples (1306 +/- 484 SE) (Figure 3.11), this was non-significant (t-test: t = 0.568; d.f. = 4; p = 0.6001), and the difference was substantially less than the roughly 300-fold difference expected.

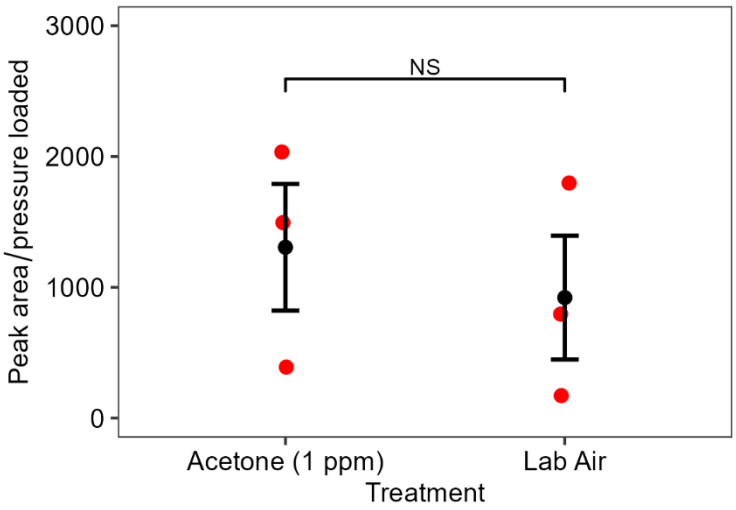


Figure 3.11: Relative acetone readings following spiking. Mean relative acetone readings were not significantly different between headspace samples from vessels spiked with 1 ppm acetone and control vessels containing lab air (t-test: t = 0.568; d.f. = 4; p = 0.6001). Error bars are +/- one standard error.

3.2.3.2 Isoprene

To investigate whether the unexpected results for acetone may have been due to some issue with the GC-MS machine specific to this volatile, a second spiking experiment was conducted for isoprene. This was done as for acetone (Figure 3.10), except this time in quadruplicate, and using a spiked concentration of 10 ppm isoprene (around 1000-fold higher than the highest estimates for atmospheric levels (Cleveland and Yavitt 1998)). Once again there was no significant difference between relative readings for the two treatments (Figure 3.12).

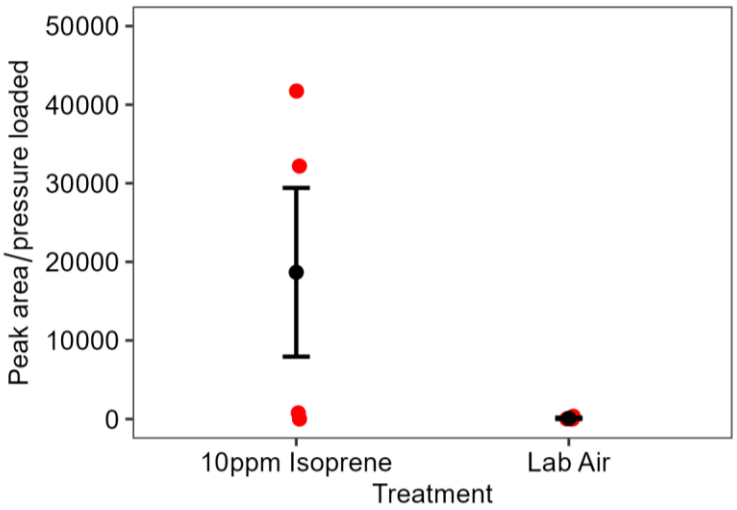


Figure 3.12: Relative isoprene readings following spiking. Mean relative isoprene readings were not significantly different between headspace samples from vessels spiked with 10 ppm isoprene and control vessels containing lab air (t-test: t = 1.731; d.f. = 6; p = 0.1342). Error bars are +/- one standard error.

3.2.3.3 Trying a different GC-MS

The above experiments raised concerns about the suitability and reliability of this equipment for the intended purpose. Preliminary experiments to assess the concentration of these volatiles in digester headspace had also generated inconsistent results (data not shown), so access was obtained to a second GC-MS, located in the Biosciences Technology Facility at the University of York (hereafter referred to as TF GC-MS), which uses the alternative Solid-Phase Microextraction (SPME) method. For this specific setup, calibration and quantification would only be possible for volatiles which can exist as liquids, meaning this equipment would not be suitable for ethylene since it is only available as a gas. Additionally, while methyl bromide can be purchased in liquid form, the hazards of working with this as a liquid compared with a bottled gas (as was used to calibrate the previous instrument) rendered further use of this volatile no longer logistically acceptable.

3.2.3.4 Testing the remaining four volatiles on the TF GC-MS

Standards for the remaining four volatiles for which constructs had been obtained or produced (isoprene, acetone, limonene and 2(methylthio)ethanol) could be obtained in liquid form. To first determine whether the TF GC-MS could detect these volatiles at all, a mixture of each of these was loaded at very high (1% by volume) concentrations onto the

machine. Analysis of the resulting traces revealed that while limonene and 2(methylthio)ethanol were readily separated from the mixture, isoprene and acetone did not appear to bind to the SPME fibre and instead eluted with air (Panel A, Figure 3.13). This was confirmed by loading individual acetone and isoprene samples at even higher concentrations: in these cases their representative ions were found to increase in proportion in this early, mixed peak.

Testing of a second GC-MS column and alternative programmes enabled the additional separation of acetone from air (Panel B, Figure 3.13), but it was still not possible to separate isoprene. Residual readings for limonene even in samples where it was not loaded revealed that there was substantial carryover of this compound between readings, presumably caused by it “sticking” to the SPME fibre.

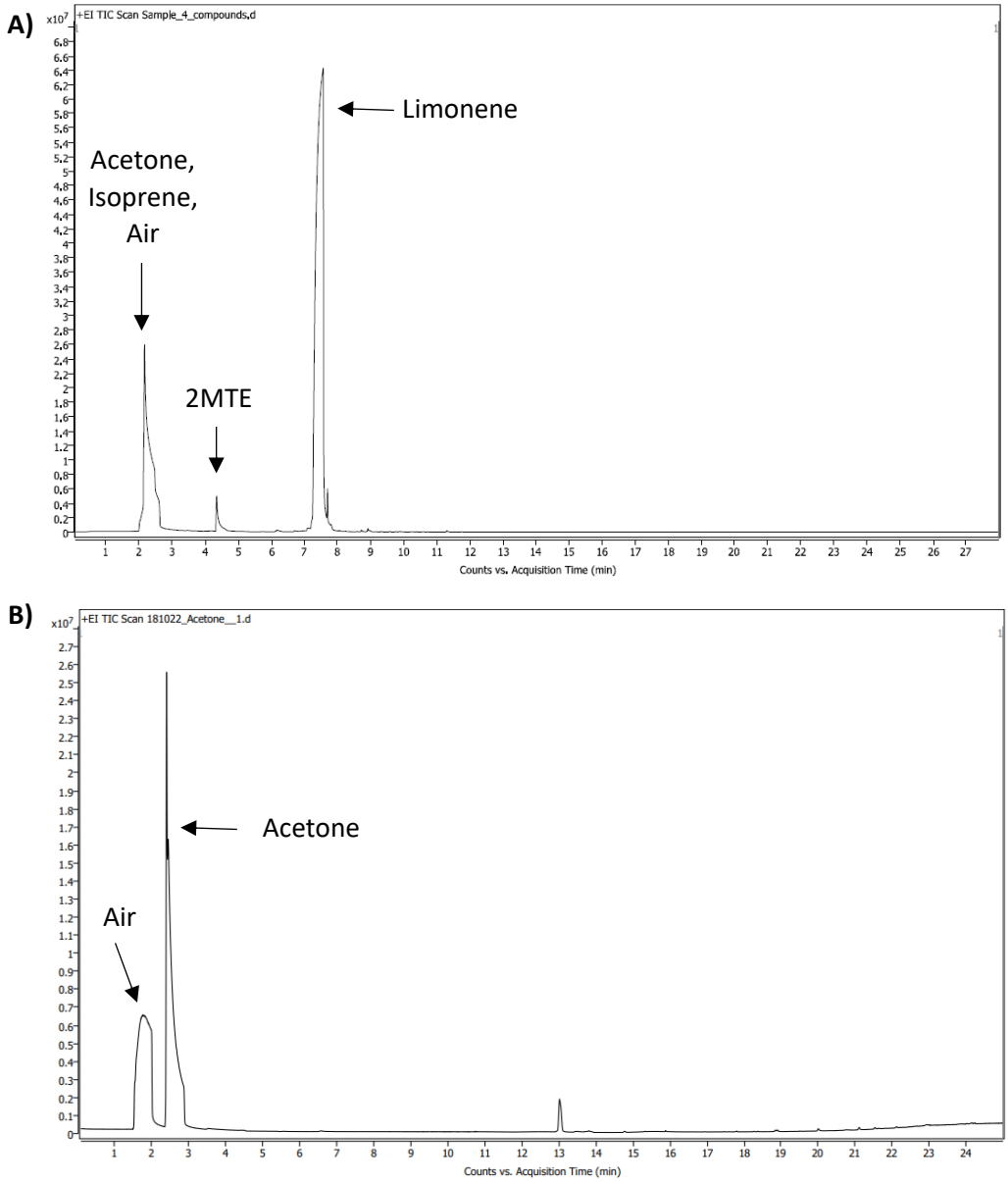


Figure 3.13: Representative traces generated on TF GC-MS.
The identity of volatiles of interest present within each major peak are indicated. **A)** Trace containing acetone, isoprene, 2(methylthio)ethanol (2MTE) and limonene in which acetone and isoprene eluted with the air. **B)** Representative trace showing acetone separated from air (this specific sample did not contain the other three compounds).

Despite these limitations it had still been possible to generate a method capable of separating these four volatiles from each other. Whether the concentration of these four volatiles could be detected at more biologically relevant levels was explored. Previous yields of each volatile generated from the plasmids obtained for their synthesis were estimated from the literature. Standards of each of the compounds were serial diluted

and run on the GC-MS to determine whether the minimum threshold of detection was within an order of magnitude of the predicted yields. In all cases except acetone, the volatiles could only be detected at 100-fold or greater concentrations than those predicted to have been generated biologically.

Table 3.3: Detectability of volatiles of interest at different dilutions on TF GC-MS.
Ticks indicate concentrations at which the indicated volatiles were clearly present in GC-MS traces, while crosses indicate concentrations at which they were not. Green boxes = approximate estimated biological yield from the respective plasmid (note no estimate was possible for 2(methylthio)ethanol since yield data for this volatile does not exist). These data are for the Inferno column. M = molar.

Dilution	Acetone	Isoprene	2MTE	Limonene
0.1 M	✓	✓	✓	✓
0.01 M	✓	X	✓	✓
0.001 M	✓	X	✓	✓
0.0001 M	X	X	X	X
0.00001 M	X	X	X	X
0.000001 M	X	X	X	X

These results suggested that it would be possible to detect acetone generated from the obtained expression system aerobically, which means this would also likely have been possible anaerobically. However, to recreate the dual volatile reporter for use in AD, two anaerobically producible volatiles were needed, and the use of the TF GC-MS had removed the other already anaerobically validated volatile (methyl bromide) from the list of candidates, leaving just one (potentially) viable candidate that could be detected.

3.2.3.5 Refocusing aims to anaerobic ethylene synthesis

Due to the difficulties encountered in detecting appropriate volatiles, and the significant amount of time that had already been invested, the decision was made not to proceed with attempting to recreate the anaerobic bioreporter. The aim was refocused solely towards the heterologous anaerobic production of ethylene from plasmids that had been produced, since this would represent a significant and biologically useful achievement. Testing the ethylene producing plasmids required an alternative detection method, since this was not an option on the TF GC-MS.

3.2.3.5.1 Ethylene detection by infrared sensor

An infrared sensor in a custom housing/sampling system had been used successfully elsewhere in this project to quantify methane gas (see section 5.2.2.1). In an attempt to extend this method for ethylene, a photoionisation detector which according to its specifications should have been able to detect ethylene at low parts per billion (ppb) levels was obtained and trialled with the existing housing/sampling system. While the sensor was found to respond to ethylene, a rapid loss of signal (later found to be caused by a leak from the back of the sensor through the housing) prohibited quantification. Attempts to modify the existing housing to prevent this leak were unsuccessful. When a second custom made 3D printed housing was produced, but signal loss was still detected, an alternative method for ethylene quantification was sought.

3.2.3.5.2 Ethylene detection by GC

An approach using a third GC instrument (without MS) was explored (hereafter referred to simply as GC). A column specifically for use with ethylene was obtained and initial testing showed a gaseous ethylene standard produced a single clear peak at the expected elution time with a relatively flat baseline. Crude dilutions showed a proportional change in peak area, confirming the origin of this peak as ethylene.

Methods for dilution of the gaseous standard were tested to enable the generation of calibration curves and to determine the limit of detection for ethylene with this assay system. Two different gas dilution methods (based on pressure and volume) were trialled (Grob 1985): dilutions by volume were found to be more reproducible so this method was selected. The lowest ethylene dilution which could be accurately created using this method was 5 parts per million (ppm). Control headspace samples containing only nitrogen (the diluent used for ethylene) often yielded a small peak at the expected elution time for ethylene, resulting in very low ethylene peak area values in most blank samples following integration. Since these values were consistently low (around 8 AU) and were generally visually distinguishable on GC traces from the lowest 5 ppm dilution (and always clearly distinguishable for the next lowest 25 ppm dilution) they were not seen as prohibitive.

A series of calibration curves spanning a range of 5 – 250 ppm ethylene were created (Figure 3.14). Different calibration curves produced from distinct, separately prepared dilution series generated very similar results, indicating that both the GC protocol and dilution method were highly reproducible, and suitable for detecting ethylene in headspace samples down to the low ppm range.

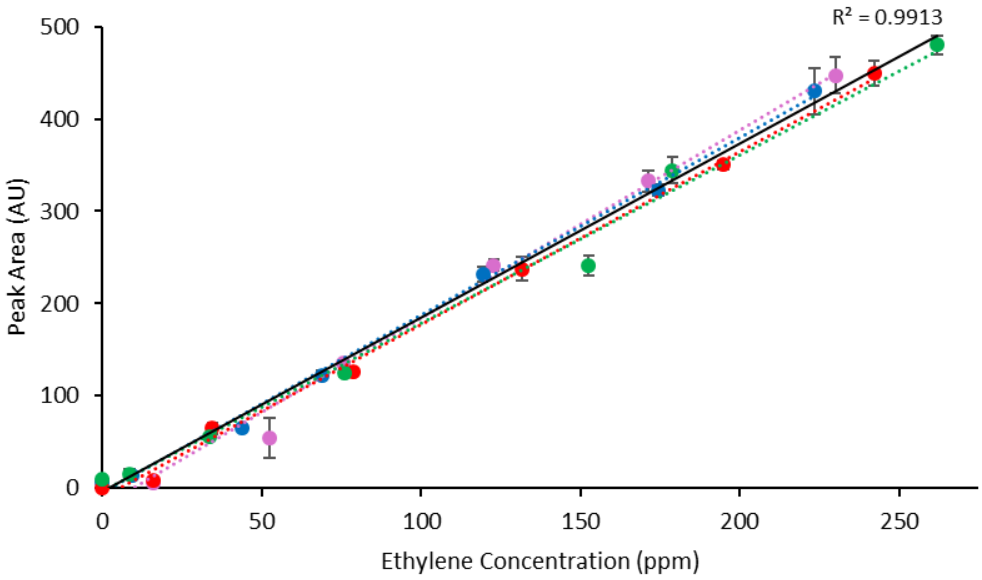


Figure 3.14: Ethylene calibrations on GC.
Four separate ethylene calibration curves (dotted coloured lines) and overall calibration curve (solid black line). Each of the four curves was prepared using an independent six-point ethylene dilution series. Points are mean peak area of presumptive ethylene peak from triplicate headspace GC runs of each dilution. Colours indicate the four different dilution series. Error bars are +/- one standard deviation. The coefficient of determination (R^2) is indicated above the overall calibration curve which incorporates all readings from all four separate dilution series.

3.2.3.6 Testing ethylene producing strains

3.2.3.6.1 Generating strains

Before the ethylene producing plasmids could be tested they needed to be transformed into their host *E. coli* strains (Table 3.2). These transformations were found to be unexpectedly challenging.

Plasmids TestC/CtrlC were transformed into *E. coli* strain MG1655 and TestB/CtrlB into *E. coli* ATCC 25922 using the same competent cell generation and transformation protocol which had been used successfully on multiple occasions previously in this chapter.

Unexpectedly, only one of these four transformations (TestB into MG1655) yielded any colonies after plate recovery, with very few colonies obtained (just six). Transformations carried out previously using this method in strain MG1655 had generated hundreds of colonies, demonstrating the cells were both viable and competent. These observations pointed to a very low transformation rate for these plasmids. All six TestC/MG1655 colonies were sequenced, one was found to have a frameshift in the *ald2* gene, but the other five had the expected plasmid sequence.

Through repeat transformations with an increased volume of cells and plasmid it was possible to obtain successful transformants for the CtrlC/MG1655 combination, which were confirmed correct through sequencing. This adjustment still yielded no colonies for either of the transformations into strain ATCC 25922 after overnight growth. Prolonged incubation to allow for a slower growth/recovery period in this strain yielded six colonies across both transformations. Further investigation revealed one of these colonies contained a cryptic *E. coli* plasmid and the others either could not be outgrown for plasmid isolation or contained no detectable plasmid, indicating these were likely *E. coli* colonies without the target plasmid(s) that had been able to grow on the plates due to antibiotic activity loss following prolonged incubation.

The literature was consulted to confirm there was no biological reason why these transformations would be expected to have failed in ATCC 25922 (e.g. presence of restriction modification systems). When no such reason could be found, a publication was identified (Alshammari *et al.* 2023) in which the authors had successfully transformed this exact strain using a chemical competence protocol (Chang *et al.* 2017). This alternative method was used to generate fresh ACTC 25922 competent cells and transform the TestB/CtrlB plasmids. Once again, no colonies were recovered for the TestB/CtrlB transformations, but a single colony recovered for the TestC plasmid (which had been included as a positive control transformation) subsequently confirmed to be correct by sequencing demonstrated that these cells were competent. Despite this, four further attempts of this transformation for both TestC and CtrlC failed to yield any colonies, indicating that for unknown reasons these plasmid/strain combinations could not be generated with the methods attempted here.

3.2.3.6.2 Ethylene production tests

While it had not been possible to generate the ATCC 25922 ethylene producing *E. coli* strains (Table 3.2), both a test and control plasmid (TestC and CtrlC) carrying the full putative ethylene pathway had been successfully transformed into commensal *E. coli* (MG1655). With a method for ethylene detection from headspace samples in hand, the decision was taken to test only these strains for ethylene synthesis.

3.2.3.6.3 Initial ethylene production test

The ethylene producing pathway heterologously expressed here is strictly anaerobic in its endogenous host (North *et al.* 2017). This could be due to repression of the pathway in non-oxygen limiting conditions (i.e. it could theoretically function in the presence of oxygen) but it could also be due to some components of the pathway being unable to operate in an oxygen atmosphere. To account for the second possibility this initial test was conducted both aerobically and anaerobically. The plasmid/strain combination TestC/MG1655, which should be able to produce ethylene, and its control (CtrlC/MG1655, which should not) were both grown with and without oxygen at two different starting cell concentrations (A – less dilute, and B – more dilute). Headspace was sampled at two different timepoints and run on the GC to check for ethylene synthesis (Figure 3.15).

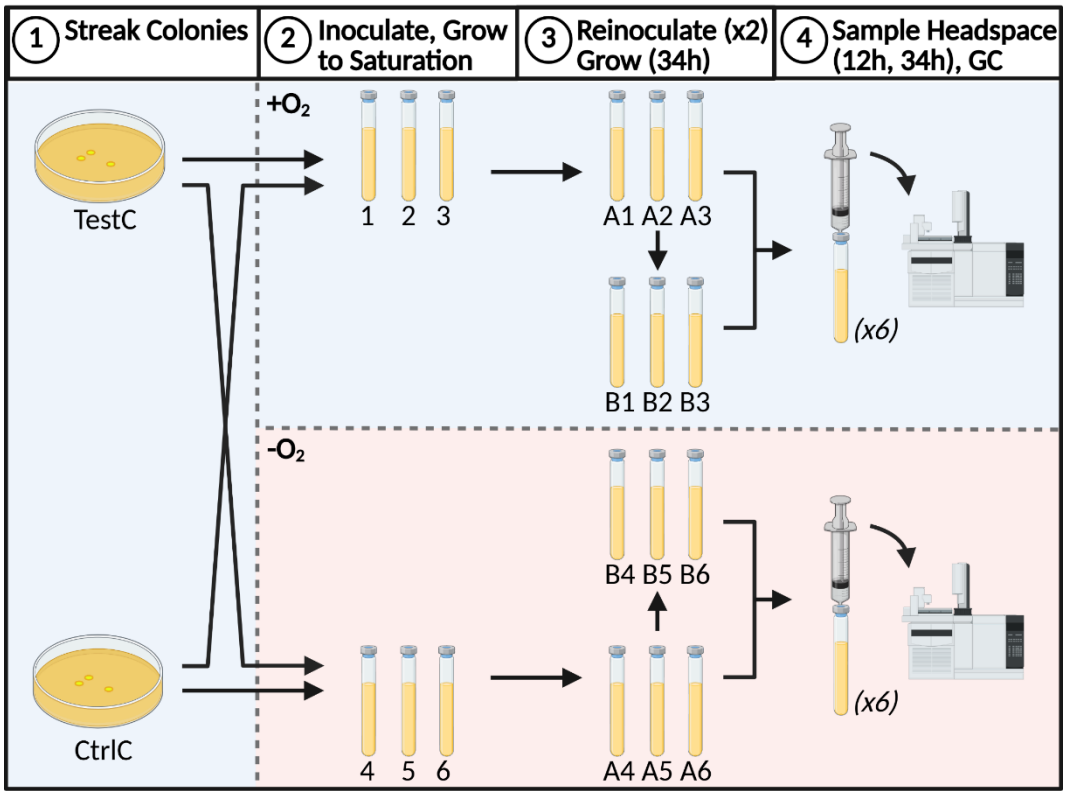


Figure 3.15: Overview of initial ethylene production test.
1) *E. coli* MG1655 harbouring the plasmids TestC and CtrlC are streaked onto plates and grown until single colonies are present. **2)** Single colonies containing each plasmid are used to inoculate triplicate starter cultures aerobically (cultures 1-3) and anaerobically (cultures 4-6), which are grown to saturation. **3)** Starter cultures are used to make a triplicate set of reinoculations with a starting OD_{600} of approx. 0.05 (A cultures), which are also further diluted to generate less concentrated B cultures, both of which have the same oxygen status as their progenitors. A and B reinoculations are incubated for 34 hours. **4)** After 12 hours headspace samples are taken from all reinoculations (A and B samples) and run on the GC to determine whether ethylene is present. These are reincubated for a further 22 hours and a second set of headspace samples is run on the GC after a total of 34 hours of incubation.

Growth dynamics of the plasmid/strain combinations were as expected (Figure 3.16): anaerobic cultures grew more slowly and to a lower final OD_{600} than aerobic, and cultures inoculated at a lower starting cell density (Figure 3.16, Panel B) entered exponential phase later than those started with more cells (Figure 3.16, Panel A). Strains harbouring the CtrlC plasmid appeared to grow slightly slower than those with TestC, which was unexpected given that CtrlC contains one fewer gene so should exert a lower metabolic burden on cells, however this difference was marginal. Proportionally larger error bars suggest there was greater variance between samples in these growth curves than had been obtained

previously for different plasmids (e.g. see Figure 3.8), likely due to less frequent OD_{600} measurements in this experiment meaning triplicate cultures were more likely to be in different growth phases at the points cell densities were recorded.

Headspace sampling performed after 12 (Figure 3.16, Panels C and D) and 34 hours of incubation (Figure 3.16, Panels E and F) revealed the presence of a small peak at the expected time for ethylene in most conditions (around 6 AU). The presence of these small peaks in most media only controls (Figure A3, Appendix), and the fact that they had been seen before when producing calibrations, suggest peaks of these low values represent background readings rather than evidence of genuine ethylene synthesis.

However, two culture conditions (TestC aerobic A and CtrlC anaerobic B) had on average larger ethylene peaks. A slight reduction in peak area by the second sampling timepoint in both cases suggested that ethylene synthesis was responsible, this had occurred at some point in the initial 12 hours of incubation.

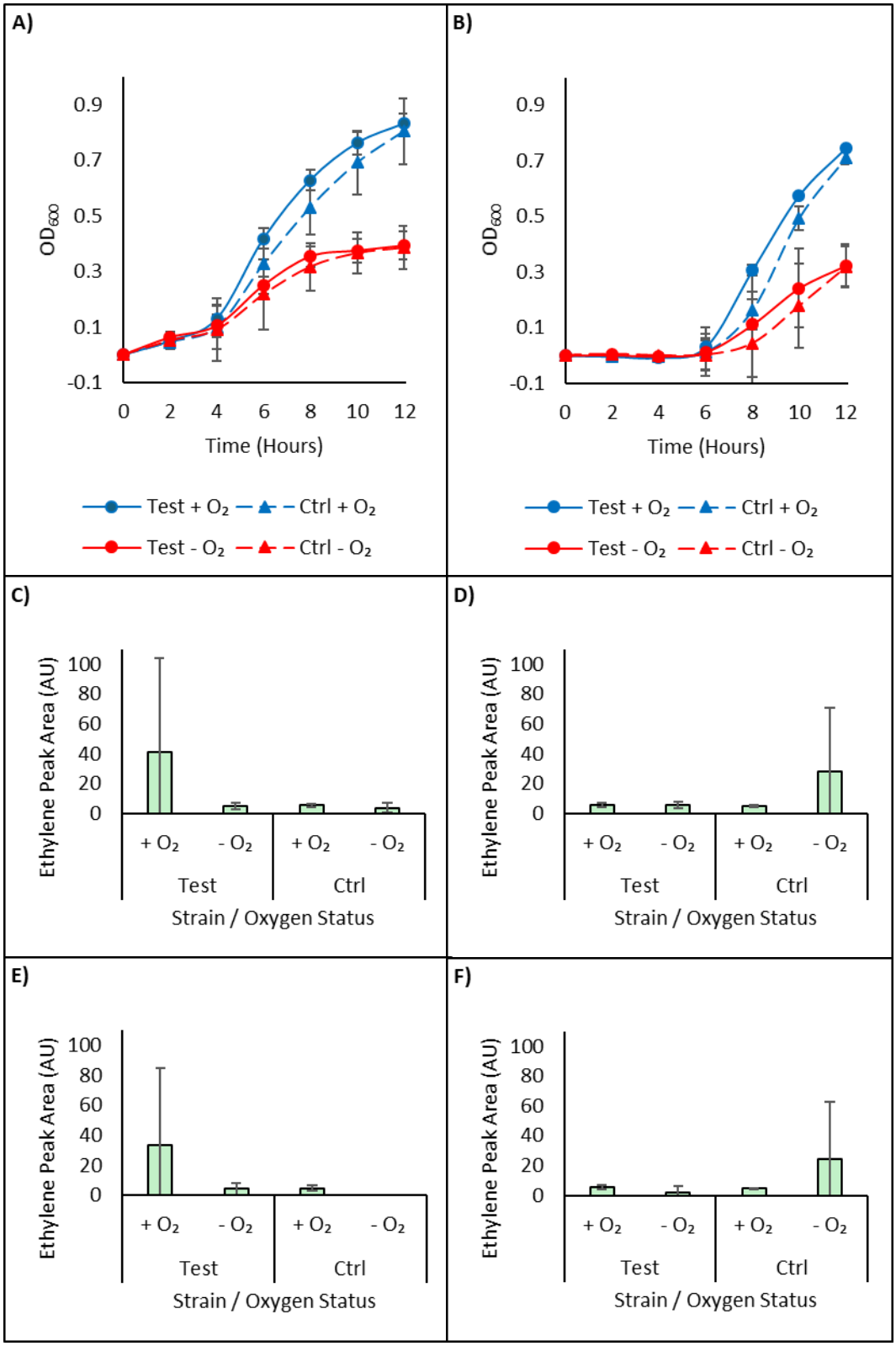


Figure 3.16: **Growth dynamics and headspace ethylene readings - initial ethylene test**. Dilutions sets A (higher starter concentration, panels A, C, D) and B (lower starter concentration panels B, D, F) of *E. coli* MG1655 cultures containing plasmids TestC and CtrlC. **A, B**) Growth curves. Point are means of triplicate cultures. Error bars are +/- one standard deviation. **C, D)** Headspace ethylene measurements after 12 hours of incubation and **E, F)** 34 hours of incubation. Bars are means of triplicate cultures. Error bars are +/- one standard deviation. Error bars extending into negative values are not shown.

Closer inspection of the GC traces (Figure 3.17) revealed two individual replicate cultures were almost entirely responsible for the comparatively higher ethylene integration values in their respective conditions (culture A3 for TestC aerobic A condition (Figure 3.17, Panel A) and culture B6 for CtrlC anaerobic B condition (Figure 3.17, Panel B)).

To verify that these larger peaks were not due to some malfunction on the GC instrument each had been run in triplicate at each sampling point (total 12 measurements) and all consistently yielded the larger peak (Figure A4, Appendix). The absence of this larger peak in controls run directly before and after some of these samples further confirmed that (at least in terms of the measurements themselves) the larger peaks in these two cultures were genuine.

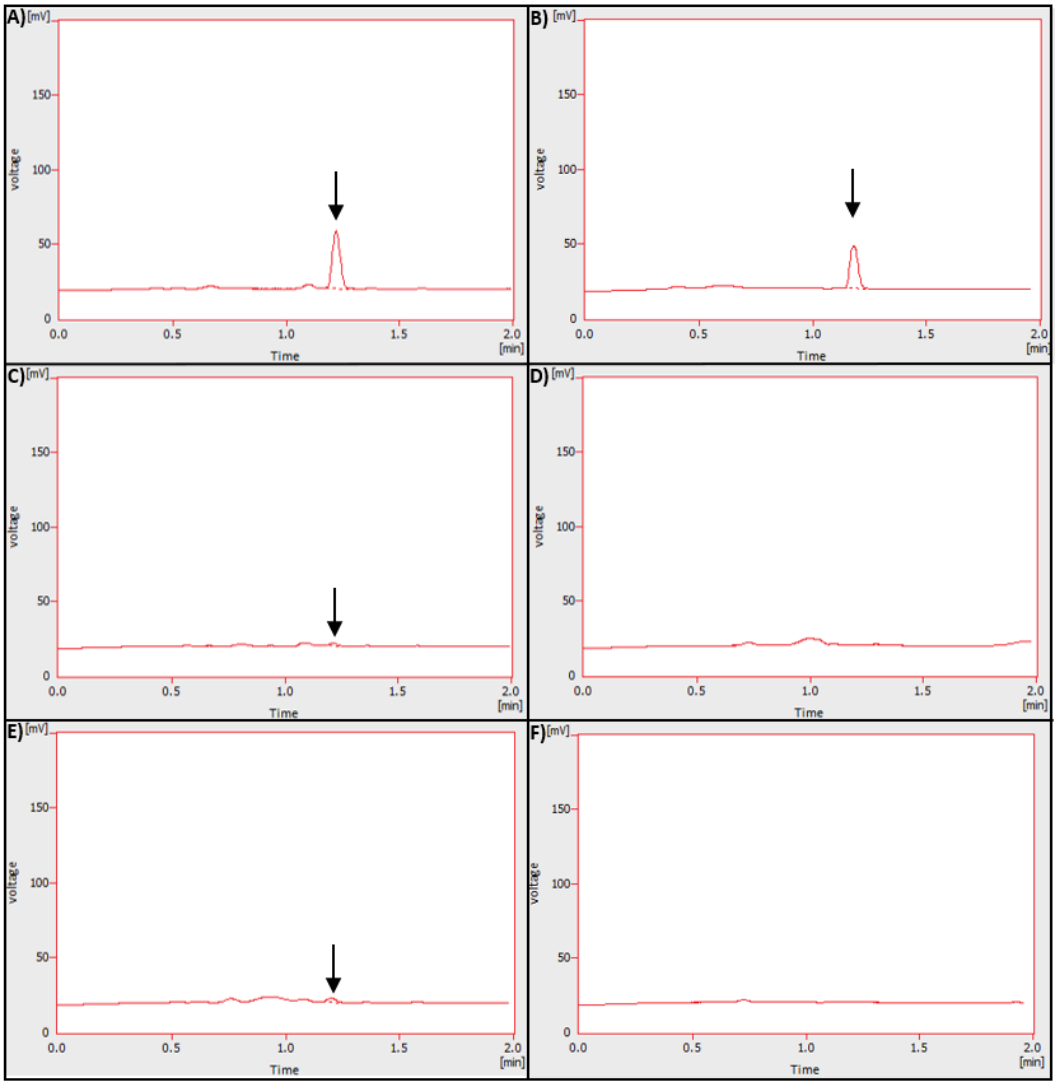


Figure 3.17: Representative ethylene traces generated on GC.
Raw GC traces from second headspace samples taken after 34 hours of incubation in initial ethylene test. **A)** TestC culture A3 containing putative ethylene peak. **B)** CtrlC culture B6 containing putative ethylene peak. **C)** TestC culture A2 - from same triplicate as culture A3 but not containing the pronounced ethylene peak. **D)** CtrlC culture B5 - from same triplicate as culture B6 but not containing an ethylene peak. **E)** Aerobic media blank control. **F)** Anaerobic media blank control. A, C, E: aerobic conditions. B, D, F: anaerobic conditions. The position of the presumptive ethylene peak is indicated with an arrow in all traces where it was detected above the baseline (and therefore integrated).

These results presented the possibility that two of the cultures tested in this initial experiment had synthesised ethylene, while the others in their respective triplicate cultures had not for unknown reasons. Glycerol stocks were generated to preserve these cultures for further testing to account for the possibility that there was something unusual about these two cultures compared with their counterparts which had enabled them to produce ethylene.

An obvious question which had to be addressed before proceeding was how one of the two cultures that had apparently generated ethylene could have done so given it contained the control plasmid which had been designed to not be able to do this.

The control plasmid CtrlC differs from the test plasmid by lacking a single gene from the ethylene synthesis pathway (*mtnK*), the absence of which should mean strains carrying this plasmid are not able to make the ethylene precursor MTR-1P (Figure 3.1) and subsequently not synthesise ethylene. However, on closer consideration having observed the above results there could be a possible explanation as to how *E. coli* cells lacking this gene may still be able to generate ethylene (see section 3.3.5).

The objective now was to replicate the putative ethylene synthesis that had been observed in the initial test (Figure 3.17). Attempts to recreate the initial results were based on the assumption that the ethylene detected previously was genuine i.e. had been biologically generated by the engineered strains. If it is assumed the results observed were due to biological ethylene synthesis, then it can be produced aerobically and by the *E. coli* strain carrying either plasmid. This meant further tests could be carried out aerobically (which is simpler than working anaerobically) with a single plasmid (TestC).

3.2.3.6.4 Varied antibiotic concentration ethylene test.

Given that these plasmids were large, and carried many (seven or eight) constitutively expressed genes, they would be expected to exert a considerable metabolic burden on their host cells and there would presumably be a strong selective pressure for the host cells to eject part or all of these constructs. It was hypothesised that the most likely explanation as to why only two of the 24 cultures tested in the initial experiment synthesised ethylene would be either that the plasmids were lost from the cells or had recombined (or a mixture of both) in 22 of 24 cases.

The first of these possibilities was investigated by repeating the headspace sampling test with different antibiotic concentrations in the growth medium, ranging from 25 µg/mL which had been used previously to 200 µg/mL. Theoretically, higher antibiotic concentrations would prevent cells which had lost the whole plasmid (and the chloramphenicol resistance gene it carries) from surviving. The experimental setup was

similar to that shown in Figure 3.15, except only MG1655 carrying TestC grown aerobically was tested. The dilution protocol for generating inoculations from starter cultures was simplified since the level of dilution did not appear to have made a difference in the initial experiment.

As would be expected, growth rates of *E. coli* cultures were lower with increasing antibiotic concentration (Figure 3.18, Panel A). However, headspace samples from reinoculations did not contain a peak above background levels at any of the tested antibiotic concentrations (Figure 3.18, Panel B). Headspace from overnight starter cultures used to generate these reinoculations were also measured for ethylene, but it was not detected above background levels in these either (data not shown). This result suggests that complete loss of the ethylene synthesis plasmid does not explain the results observed in the initial test, since these cultures should not have been able to survive without the antibiotic resistance gene it carries, particularly at higher antibiotic concentrations.

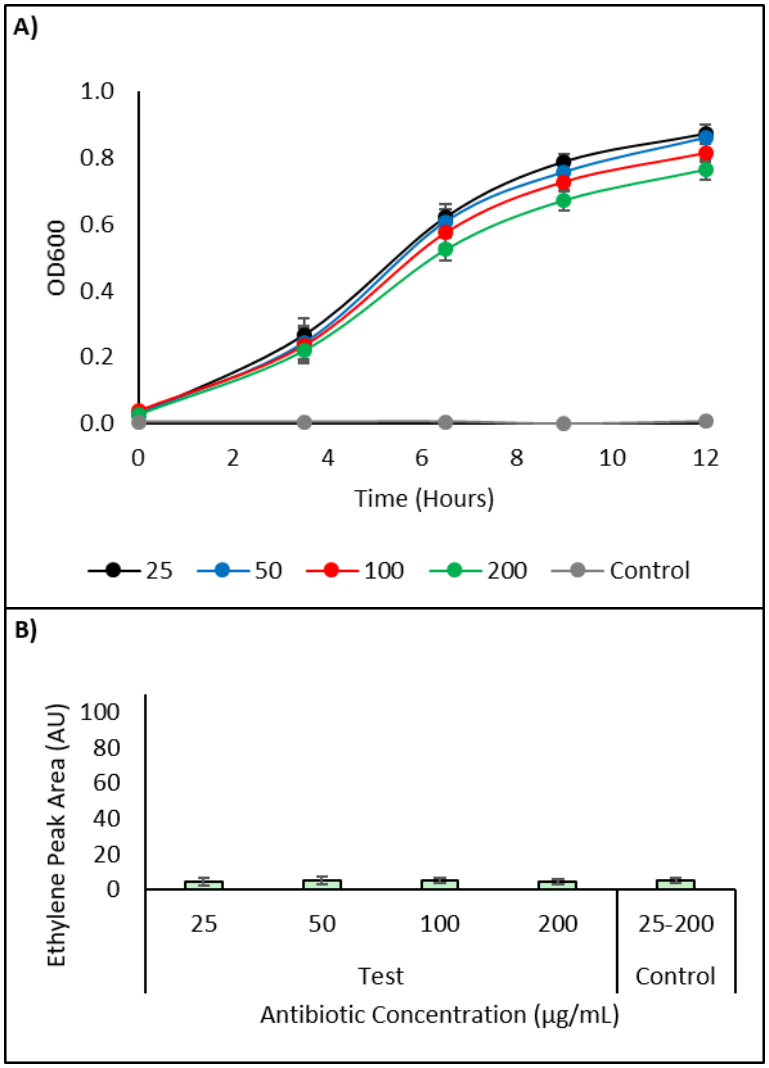


Figure 3.18: Varied antibiotic concentration ethylene test.
E. coli MG1655 cultures containing plasmid TestC at different chloramphenicol concentrations. **A)** Growth curves. Points are means of 10 replicate E. coli cultures except for the control (grey line) which is a mean of four single mock inoculated cultures at each antibiotic concentration (µg/mL). Error bars are +/- one standard deviation. **B)** Headspace ethylene measurements after 12 hours of incubation. Bars are means of 10 *E. coli* cultures, except for the control which is a mean of four readings from a single mock inoculated culture at each antibiotic concentration (µg/mL). Error bars are +/- one standard deviation.

3.2.3.6.5 DH5α ethylene test

To test the possibility that the results observed in the initial ethylene test may have been due to plasmid recombination, which would theoretically enable the strains to retain the antibiotic resistance gene but not the ethylene synthesis components of the plasmids, the headspace sampling test was carried out using the TestC plasmid in the DH5α *E. coli* strain.

Strain DH5α is used mainly for plasmid manipulation and has been engineered to be recombinase deficient, and would not be expected to be able to recombine as easily as the strain used up to this point for these experiments (MG1655). DH5α is not typically used for protein expression and expression of the pathway components in this strain would be expected to be very low, however given that even low expression might be sufficient to generate detectable levels of ethylene, it seemed sensible to try. For this test, in addition to single colonies of DH5α carrying the TestC plasmid, an additional culture derived from multiple colonies was included to address the possibility that stochastic differences between colonies may have been the explanation for the differing putative ethylene production observed in the initial experiment.

Growth dynamics for these cultures was as expected (Figure 3.19, Panel A), but ethylene was not detected above background levels in any of the treatments indicating no ethylene synthesis had occurred, suggesting recombination of plasmids had not occurred given that this would be extremely unlikely in DH5α.

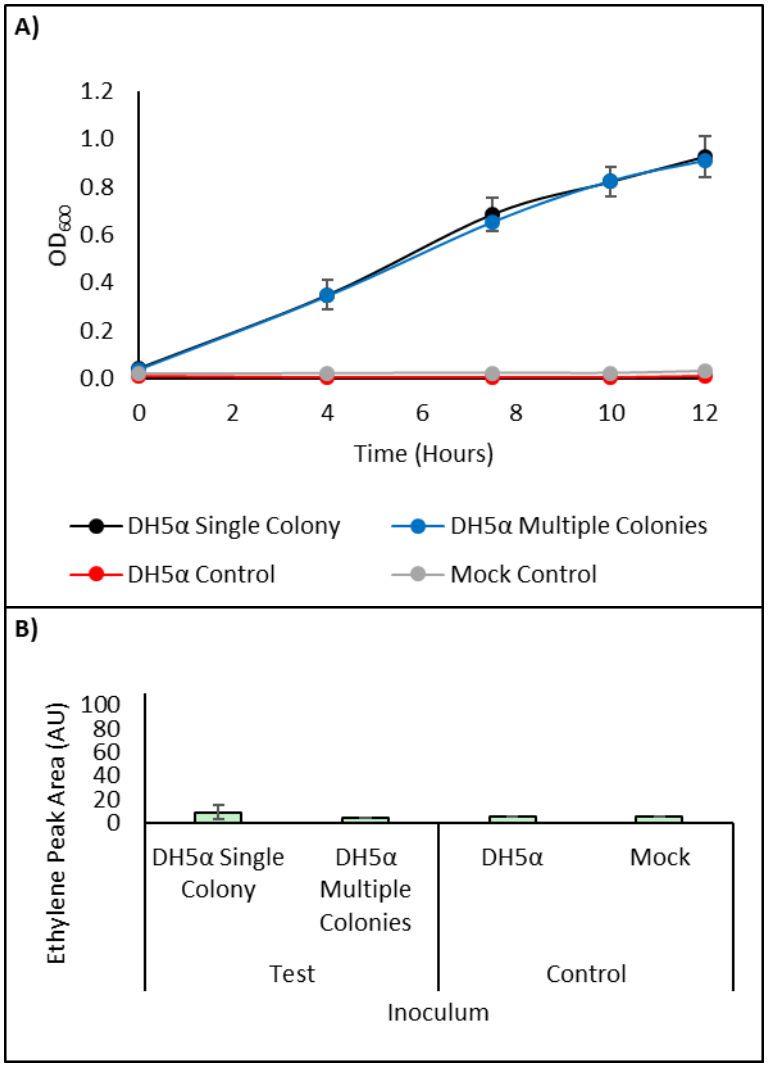


Figure 3.19: DH5α ethylene test.
E. coli DH5α cultures containing plasmid TestC. **A)** Growth curves. Points for single colony cultures (black line) are means of four replicates, all other curves are single cultures. Error bars are +/- one standard deviation. **B)** Headspace ethylene measurements after 12 hours of incubation. The bar for single colony cultures is a mean of four replicates, all other bars are derived from a single culture. Error bars are +/- one standard deviation.

The conclusion that recombination was not the cause of the results in the initial ethylene expression experiment was further confirmed by later sequencing of plasmids obtained from pelleted material of various strains tested for ethylene synthesis which had been retained throughout this process (Table 3.4). While evidence of recombination events which would render the plasmid non-functional were detected in three cases in the MG1655 strain, in the majority of cases all functional elements of the plasmid had

remained intact despite no ethylene having been generated from any of the strains from which they were derived.

Table 3.4: Plasmid status in *E. coli* cultures following ethylene synthesis tests.
Sequence analysis summary of plasmids isolated from *E. coli* cultures which had previously been tested for ethylene synthesis. Plasmids were deemed to be unable to support ethylene synthesis (expected non-functional) if one of more of the ethylene pathway genes had been lost due to recombination. Plasmids in which all pathway components remained intact were deemed expected functional.

Plasmid	*E. coli* Strain	Colonies Sequenced	Expected Functional	Expected Non-functional
TestC	MG1655	8	5	3
TestC	DH5α	4	4	0
CtrlC	MG1655	4	4	0

3.2.3.6.6 Repeat testing of putative ethylene producing cultures

Having ruled out the two most likely explanations as to why only 2/24 of the cultures initially tested for ethylene production were able to produce it, the possibility that there was something unusual about the *E. coli* present in these two cultures was investigated.

While it seemed unlikely, it was possible that (for example) some kind of mutation within the genome of the strains themselves permissive for ethylene production from the plasmids has taken place.

To test this the glycerol stocks of cultures TestC A3 and CtrlC B6 which had been generated in the initial experiment were revived and retested for ethylene production similarly to previous experiments. Strain growth was as expected (Figure 3.20, Panel A), but no ethylene was detected above background levels in headspace samples from reinoculated cultures (Figure 3.20, Panel B) or the overnight starter cultures used to generate them (data not shown).

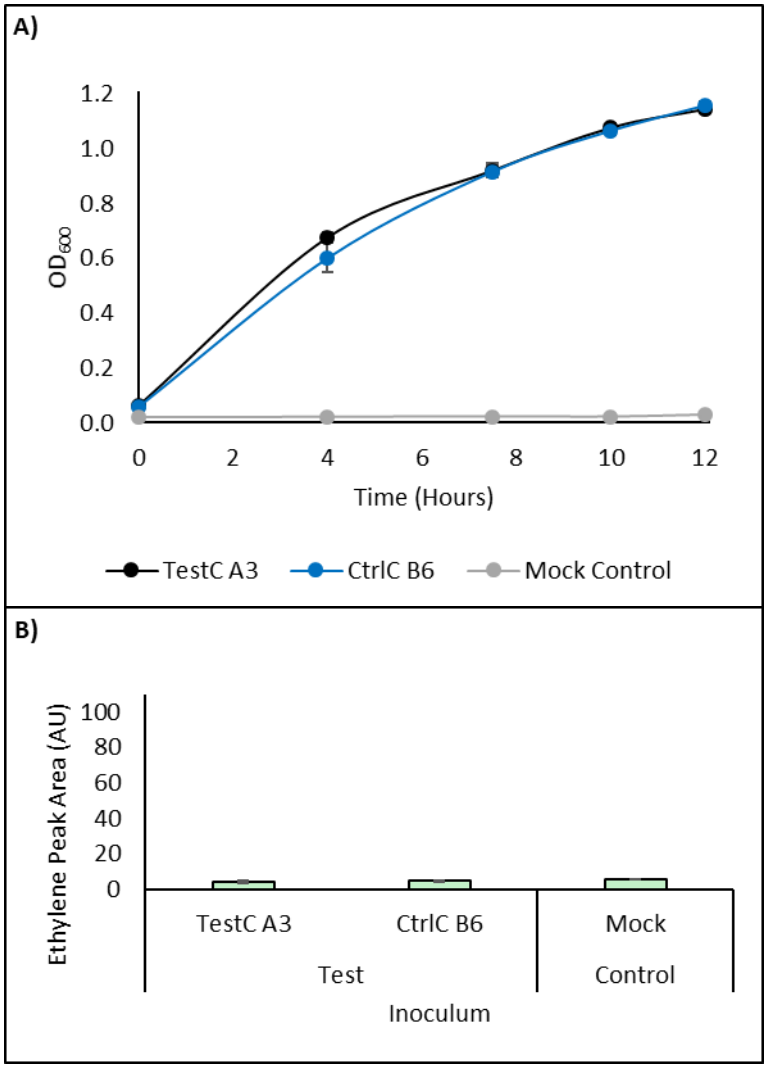


Figure 3.20: Repeat of initial culture ethylene test.
E. coli MG1655 cultures containing plasmid TestC derived from the two cultures containing a putative ethylene peak from the initial test. **A)** Growth curves. Points for A3 and B6 cultures are means of four replicates, the control is a single culture. Error bars are +/- one standard deviation. **B)** Headspace ethylene measurements after 12 hours of incubation. The bar for A3 and B6 cultures is a mean of four replicates, the control is a single culture. Error bars are +/- one standard deviation.

3.3 Discussion

3.3.1 Aims recap

The aim of this chapter was to repurpose an existing quantitative whole cell volatile bioreporter (Cheng *et al.* 2018) for use in AD. The key step to achieving this has been to identify an expression system capable of producing a second detectable volatile anaerobically. When it became apparent this goal would not be achieved, work was refocused to attempt to anaerobically generate one of the candidate volatiles (ethylene).

3.3.2 Challenges encountered

The greatest difficulty encountered in testing additional candidate volatiles was not in generating/obtaining the required constructs, but rather in detecting the volatiles themselves. Early issues with the GC-MS instrument available meant a different instrument had to be used. While it was possible to detect some of the selected volatiles using a second GC-MS, the configuration meant it was not suitable for ethylene or methyl bromide, and was unable to detect three of the four remaining candidate volatiles at biologically relevant levels. Different GC-MS setups are configured to detect different classes of volatiles, so it is perhaps unsurprising that these difficulties were encountered when moving from the first machine (to which the choice of candidate volatiles had been specifically tailored) to the second (to which it had not).

3.3.3 Chapter achievements

Given the difficulties explained above, the focus of this portion of the project was shifted to attempting to anaerobically generate ethylene heterologously in *E. coli*, using a recently discovered pathway (North *et al.* 2017; North *et al.* 2020a). Ethylene is the most in demand precursor in the chemical industry (Leite and Hallenbeck 2014) and the ability to biosynthesise it anaerobically would be useful for two reasons. Firstly, anaerobic systems are generally more economical than aerobic alternatives since exogenous oxygen does not need to be supplied, reducing energy costs (Kulakowski *et al.* 2023). Secondly, ethylene gas is explosive in an oxygen containing atmosphere (Y. Wang *et al.* 2023), which means anaerobic synthesis would logistically be much more desirable.

A series of constructs which, based on literature available at the outset of this project for this newly discovered pathway (North *et al.* 2017; North *et al.* 2020a), should contain all the components necessary for ethylene synthesis anaerobically were produced. A method capable of reliably detecting ethylene in small (1 mL) headspace samples down to low ppm ranges was also developed using a third GC instrument. On initial testing of the larger ethylene plasmids (TestC and CtrlC, Figure 3.5) containing the full ethylene synthesis pathway in commensal *E. coli*, there was some evidence of ethylene synthesis in two of the 24 tested cultures (Figure 3.15, Figure 3.17). Multiple attempts to recreate this result were unsuccessful.

3.3.4 Was the ethylene putatively detected genuine?

There are many possible explanations as to why it was not possible to repeat this result, but perhaps the most important to consider is whether the observed ethylene peaks were erroneous readings of some kind. Several pieces of evidence strongly suggest the readings were genuine. Firstly, this larger ethylene peak was always present in repeat loadings of headspace samples from the two cultures in which it occurred across two different timepoints (12 samples total), which strongly suggests that whatever was responsible for this peak was contained within the headspace of these samples and these peaks were not, for example, caused by a pocket of ethylene trapped in the GC from a previous ethylene standard.

A second important consideration is whether these two culture tubes had somehow become contaminated with ethylene standard. Again, this explanation seems unlikely given that ethylene standards were prepared in different glassware both spatially and temporally separately from the ethylene culture tests. Additionally, the ethylene standard used here was a gas. While contamination with liquid standards is easily done, this is less likely with a gaseous standard given that it has to be moved from vessel to vessel in gastight syringes and is rapidly lost in any non-airtight vessel. Finally, if contamination were an issue it seems unlikely that it would only have occurred in two of the approximately 100 cultures from which headspace samples were assayed for ethylene in this work.

3.3.5 Ethylene production from the control plasmid

If these two cases of ethylene synthesis were genuine, some explanation has to be provided as to how the control plasmid was able to produce it here. The only difference between the plasmids TestC and CtrlC is that the control plasmid lacks the gene *mtnK*, which codes for a kinase in some *E. coli* ExPEC strains such as ATCC 25922 (North *et al.* 2020b). This enzyme carries out the first step in the pathway that generates 2(methylthio)acetaldehyde (the precursor chemical for ethylene synthesis in organisms such as *R. rubrum*) by converting MTR to MTR-1P, (Figure 3.1) (North *et al.* 2020b). MTR-1P is subsequently converted to 2(methylthio)acetaldehyde through the sequential action of an isomerase (*mtnA*) and aldolase (*ald2*) (North *et al.* 2020b), genes for which were both present on the TestC plasmid.

Promiscuous activity of kinases has been reported in *E. coli* (Campbell *et al.* 2023), so one possibility is that other enzymes present in *E. coli* were able to carry out the reaction normally catalysed by *mtnK*. Deletion of the *mtnK* gene was found to arrest the pathway at MTR in strain ATCC 25922 (North *et al.* 2020b), which could cast doubt on this explanation. However, the strain used in this work (MG1655) is a commensal rather than ExPEC. To the author's knowledge, expression of *mtnA* and *ald2* without *mtnK* has not been reported in commensal *E. coli*, so it is possible that some other gene found in strain MG1655 but not in ATCC 25922 was able to compensate for *mtnK*. Other enzymes capable of performing this role are known, for example the phosphorylase MntP found in *R. rubrum* (North *et al.* 2017).

3.3.6 Why could the putative ethylene synthesis result not be reproduced?

If the ethylene thought to have been generated in this work was genuine, and was made in both TestC and CtrlC containing strains, why could the result not be repeated? Subsequent tests ruled out the possibility that plasmid loss and/or recombination was the cause. In the two cases where ethylene was thought to have possibly been generated, it appears this occurred at some point in the first 12 hours of incubation, since headspace readings had not increased by the second, later timepoint (Figure 3.16). One possible explanation is that these two cultures entered some kind of permissive state for ethylene synthesis for a short time window during this period which was not achieved in the other cultures, perhaps due to stochastic differences between them, for example their relative

position in the incubator making the temperature slightly different from the other cultures at some point during the incubation.

3.3.7 Possible future work for this arm of project

Immediate further work should involve additional attempts to reproduce the ethylene synthesis putatively observed here. Options which involve no, or minimal modification of the existing assay system/plasmids should be attempted first as they would be relatively quick to do. Trialling a range of lower temperatures during the outgrowth phase could lead to better expression of the ethylene pathway components as this is associated with reduced rates of protein aggregation and proteolytic degradation in *E. coli* (Francis and Page 2010). Performing the ethylene assay in freshly transformed cells as opposed to from glycerol stocks may also support better protein expression (Bhadra *et al.* 2022).

Another option would be to conduct the assay in minimal, sulphate free medium. In the endogenous host, the primary function of the pathway heterologously expressed here is for sulphur salvage: it is repressed when sulphate is readily available (North *et al.* 2017; North *et al.* 2020a). The LB medium used for the ethylene assays in this work contains sulphates (Kertesz 2000). This had not been anticipated to be a problem as the pathway components were expressed from constitutive promoters, so should have been free from regulation. However, these genes were identified in the endogenous host through a combination of proteomic and mutagenesis screens (North *et al.* 2020a) so there could be other unidentified factors involved in the pathway or its regulation that were not detected, which if also present in *E. coli* could prevent expression in non-sulphate limiting conditions. Another possible benefit to testing in minimal medium is that the metabolism of *E. coli* differs relative to in rich, complex medium like LB (Kim and Kim 2017) which could shift cells to a state more permissive for ethylene generation, for example by increasing the pools of precursor chemicals such as SAM (North *et al.* 2020a). This approach could be further complemented by more systematic consideration of cellular metabolism, for example by integrating information from *E. coli* metabolic networks and models (Orth *et al.* 2011), and using techniques such as flux balance analysis to determine how the strain and/or medium could be tuned to better support the heterologously expressed ethylene pathway (Swayambhu *et al.* 2020).

If these troubleshooting steps were unsuccessful, more time consuming options involving redesign of the constructs and generation of new strains would be needed. The pathway components could be placed under the control of inducible promoters, which can be advantageous as they decouple cell growth from recombinant protein expression (Studier 2005), which is particularly useful in the case of highly toxic proteins (Saïda *et al.* 2006). Protein tags could also be included in the ethylene pathway enzymes to ensure all of the components are being correctly expressed. This would likely be preferable over a qPCR based approach if some kind of permissive phase were the cause of the result observed here as it would be easy to miss the window in which the ethylene pathway genes were expressed at the mRNA level given the inherent chemical instability of RNA (Kornienko *et al.* 2024). Another option would be to express the ethylene pathway genes from the bacterial chromosome, as this generally confers more stable expression than plasmid systems (Ou *et al.* 2018).

Another obvious next step would be to continue to troubleshoot the transformation of the smaller ethylene pathway plasmids produced here (TestB and CtrlB) into *E. coli* ATCC 25922, such that these plasmids could be tested for their ability to produce ethylene. These plasmids contain three fewer genes than their larger counterparts since these genes are endogenously produced in *E. coli* strain ATCC 25922. It had been envisaged that these plasmids would have been more likely to work by virtue of their smaller size and the requirement for three fewer genes to need to be correctly expressed in an exogenous host. Despite multiple attempts using different competence/transformation protocols it was not possible to achieve this transformation. This is in itself tantalising since one possible explanation for this is that the plasmids were toxic to the cells, which might suggest they were capable of driving high levels of protein expression (James *et al.* 2021) for the ethylene pathway genes. A sensible place to proceed with this would be to attempt an electroporation protocol since this is generally associated with higher transformation rates than chemical methods in bacteria (Ren *et al.* 2019). With more time some or all of the above suggestions would have been attempted, but unfortunately due to time constraints this was not possible.

Other plasmids designed to produce 2(methylthio)ethanol via the precursor product 2-(methylthio)acetaldehyde were generated but not tested in this work due to time

constraints. Between the conception of this work and the writing of this thesis the successful expression of these genes in commensal *E. coli* using a design principle similar to the one proposed here has been achieved (Huening Katherine *et al.* 2024), although whether 2(methylthio)ethanol and/or 2-(methylthio)acetaldehyde could be synthesised in this strain was not covered (Huening Katherine *et al.* 2024). To the author's knowledge exogenous anaerobic synthesis of ethylene using this pathway has still not been reported. The work reported here represents a sensible starting point based on the available literature and could inform further work in this area. The Golden Gate system used to produce these constructs could be used to test additional pathway related genes and combinations in further work.

3.3.8 Next steps

The aim of this chapter (i.e. to generate an anaerobic bioreporter) was extremely ambitious and while this was ultimately not achieved, some interesting results were obtained when the focus was shifted to attempt to anaerobically generate ethylene.

The generation of a bacterial strain capable of performing some useful function anaerobically (as was the focus of this chapter) is just one part of the challenge in terms of developing biological methods for monitoring and/or manipulating AD communities, and there are other avenues which could be explored.

Chapter 4: Chassis Organisms

4.1 Introduction

4.1.1 Context of work

In Chapter 3 attempts were made to create a whole cell bioreporter, a specific biological tool for use in AD. Crucially, an engineered microbe such as a biosensor can only be useful if it can survive in the environment it is to be used in. This is mainly dictated by the host or "chassis" organism which the introduced genetic components are housed in (Adams 2016; Del Valle *et al.* 2020a), and will form the focus of this chapter.

4.1.2 Desirable features of AD chassis organisms

To be useful in AD, a chassis organism would not only have to be able to successfully invade an existing mixed microbial community but also then persist within it to perform its programmed function (Innard and Chong 2022). An anaerobic digester presents a potential challenge for any introduced organism due to possible variations in abiotic conditions such as pH or temperature (Giangeri *et al.* 2024), as well as biotic factors (i.e. competition) (Hibbing *et al.* 2010; Jiang *et al.* 2021b). This challenge is likely to be even greater for an engineered organism due to the additional metabolic burden of expressing any exogenously introduced components (Roell *et al.* 2019). Such an organism would ideally have already available genetic tools as developing these for poorly characterised organisms from scratch is extremely laborious (Wang *et al.* 2019). Finally, an ideal organism would be facultatively anaerobic (i.e. capable of growth with and without oxygen) (André *et al.* 2021) to make it easier to work with.

4.1.3 Potential strategies to increase robustness of chassis organisms in AD

There are several approaches that could be used to generate more robust chassis for use in AD (Figure 4.1). Perhaps the simplest possibility would be to isolate an endogenous microorganism from an existing anaerobic digester, engineer it to perform some desired function, then reintroduce it (Figure 4.1, Panel 1). Another approach which also makes use of existing endogenous species would be to transform them in situ (Figure 4.1, Panel 2) by introducing a lab strain that conjugates a genetic cargo carrying the desired functionality into existing community members (Wang *et al.* 2019; Rubin *et al.* 2022). An

alternative approach which does not rely on AD endogenous organisms would be to adapt a laboratory chassis to be more resilient in an AD environment using directed evolution (McDonald 2019) (Figure 4.1, Panel 3). Another possibility would be to create an environment that is more favourable to an existing laboratory chassis within a digester, either through physical separation (Figure 4.1, Panel 4) or by metabolic niche creation, as has been demonstrated in an anaerobic gut system (Shepherd *et al.* 2018). For this chapter the decision was taken to focus on isolation of endogenous organism(s) (approach 1) and directed evolution (approach 3).

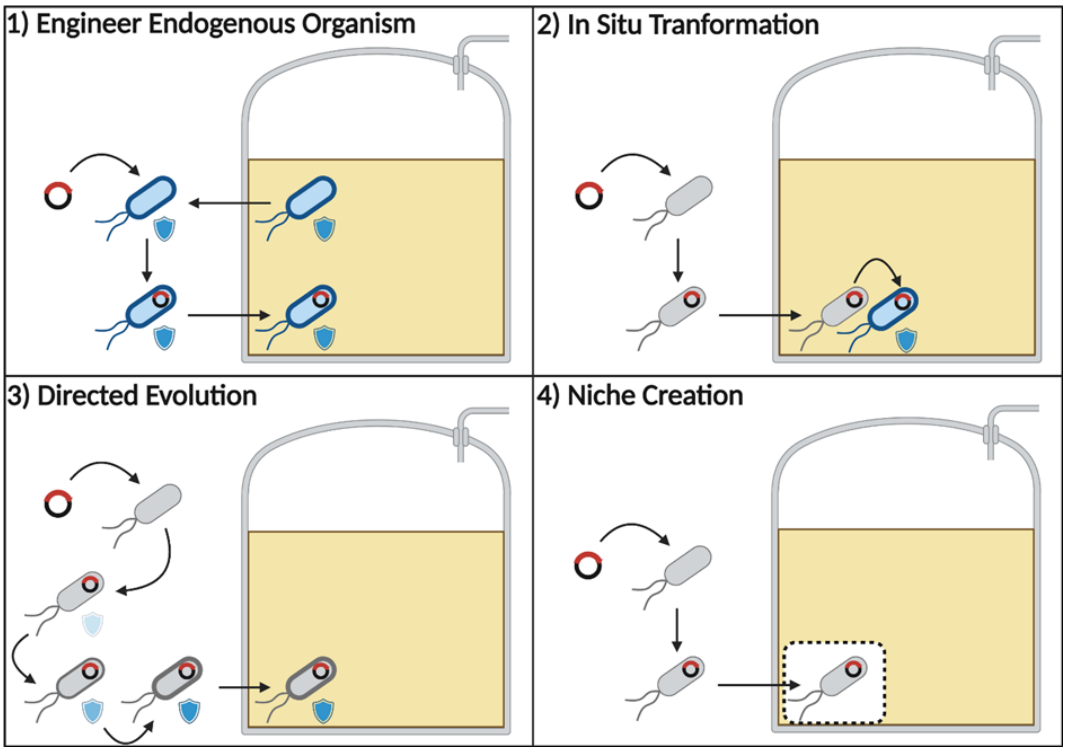


Figure 4.1: Possibilities for obtaining or generating more robust AD chassis organisms. **1)** An organism endogenous to a digester could be extracted and genetically engineered. **2)** A laboratory chassis could be used to deliver a genetic construct to endogenous members of an AD community. **3)** A laboratory chassis could be evolved to increase its fitness with respect to some rationally selected feature(s) of a digester. **4)** A laboratory strain could be in some way partially protected from the harsh conditions of the digester, for example with a physical barrier that prevents direct competition with endogenous organisms.

4.1.4 Directed evolution

Directed evolution is a well established method through which natural selection can be leveraged to increase the fitness of a microorganism with respect to a condition of interest

(Van den Bergh *et al.* 2018). A test population is exposed to an environmental condition (e.g. a toxic product) creating a selective pressure whereby adaptations or mutations that confer an advantage in that condition are selected (e.g. tolerance to toxic product), and the population becomes fitter in that condition (McDonald 2019). The evolved populations are then compared to controls in a competition assay to determine whether the process has led to fitness changes, usually by measuring some proxy of fitness (for example growth) (McDonald 2019). While the ancestral (i.e. progenitor) population can serve as a control, a dedicated control which has been evolved alongside the test population but in the absence of the condition of interest (e.g. no toxic product) can provide more useful information as it controls for unintended adaptations to experimental conditions other than the condition of interest (Santos *et al.* 2024).

The standard competition assay is direct competition, where the evolved and control strain are competed in the condition of interest (e.g. same culture tube). An essential prerequisite is that it must be possible to distinguish the strains from each other, for example with a marker gene (McDonald 2019). When using separate strains, it is also important that any differences between them such as marker genes are selectively neutral to ensure both start equally fit in the test condition (de Visser and Lenski 2002; Boynton *et al.* 2017). An alternative option is to use an approach where strains are grown separately (e.g. different culture tubes) in the same conditions, and fitness differences are inferred by comparing differences in growth rates (Kinsler *et al.* 2023). With this comparative growth assay approach the strains do not necessarily need to be distinguishable.

4.1.5 Isolating endogenous organisms from an AD community

Since an endogenous AD organism would already be adapted to the AD environment, and have undergone minimal laboratory cultivation and manipulation (by comparison to a laboratory chassis), it would be expected to be more competitive in the digester environment. An approach based on searching AD communities for “wild” alternatives of well characterised laboratory strains would come with the advantage that it should be possible to retrofit genetic tools from the laboratory strains into their wild relatives, making future genetic manipulation efforts in the isolated organisms easier. This approach also simultaneously provides an inbuilt benchmark organism against which to

compare the fitness of the isolated wild strain (in the form of its laboratory relative). Once desirable candidate organisms have been selected, it should be possible to isolate them from a mixed community by creating an environment that favours those organism(s) by plating.

4.1.6 Aims

The aim of this chapter is to investigate possible methods for obtaining or generating new chassis organisms suitable for an AD environment. This will involve:

1. Attempting to isolate organisms from sludge that have close relatives with already available genetic tools.
2. Demonstrating that a laboratory microorganism can be evolved to be fitter in some parameter relevant to AD.

4.2 Results

4.2.1 Chassis organisms: directed evolution

One approach selected for obtaining more robust chassis organisms was to evolve a laboratory microorganism to be fitter in some parameter relevant to AD. The anaerobic environment itself (i.e. the absence of oxygen) was selected as the factor since it is fundamental to all AD systems.

4.2.1.1 Obtaining suitable organisms

A laboratory microorganism needed to be obtained to act as a starting point for directed evolution. *E. coli* was selected because it is the best-established bacterial model (Meier-Kolthoff *et al.* 2014) and can be grown both aerobically and anaerobically (Conway and Cohen 2015). The strain needed to contain a selectable marker suitable for it to be distinguished from and isolated out of a mixed microbial AD community following evolution experiments.

A pair of *E. coli* strains which met this requirement were obtained. Strain MG1655_rLP5(nth-ydgR) constitutively expresses red fluorescent protein (RFP) and a chloramphenicol resistance gene from the chromosome (Urtecho *et al.* 2019) (hereafter simply referred to as "RFP expressing strain"). Strain MG1655λatt-kmgfp expresses green fluorescent protein (GFP) and a kanamycin resistance gene also from the chromosome (Da Re *et al.* 2007) (hereafter referred to as "GFP expressing strain"). Chromosomal expression was selected as it is associated with higher stability than plasmid based expression (Saleski *et al.* 2021), which meant these strains should be usable without addition of antibiotics and so could be cultured within an AD community (where the majority of organisms would be expected to be susceptible). These antibiotic resistance genes were chosen as the mechanisms of action of both is internal to the cell (Piñeiro *et al.* 1994; Zelazny and Bibi 1996) (i.e. the product cannot be excreted and "shared" with other organisms not carrying the genes) as this could confound evolution experiments.

The selection of two strains provided contingency in case one did not perform as expected, and the combination of different visual and antibiotic markers was chosen so they could in principle be used together in direct competition experiments. Both strains

are based in the *E. coli* MG1655 background, so MG1655 without these markers was used as the control organism (hereafter referred to as "WT").

For the planned evolution experiments to work the strains needed to be visually distinguishable on plates from WT cells, and also ideally from each other. To test this, all three strains were grown aerobically in liquid culture, mixed in all combinations at equal ratios and plated. As expected colonies were indistinguishable under normal light (Figure 4.2, Panels A-C), but fluorescent imaging enabled the GFP and RFP expressing strains to be distinguished from each other (Figure 4.2, Panel D) and from WT cells (Figure 4.2, Panel E-F).

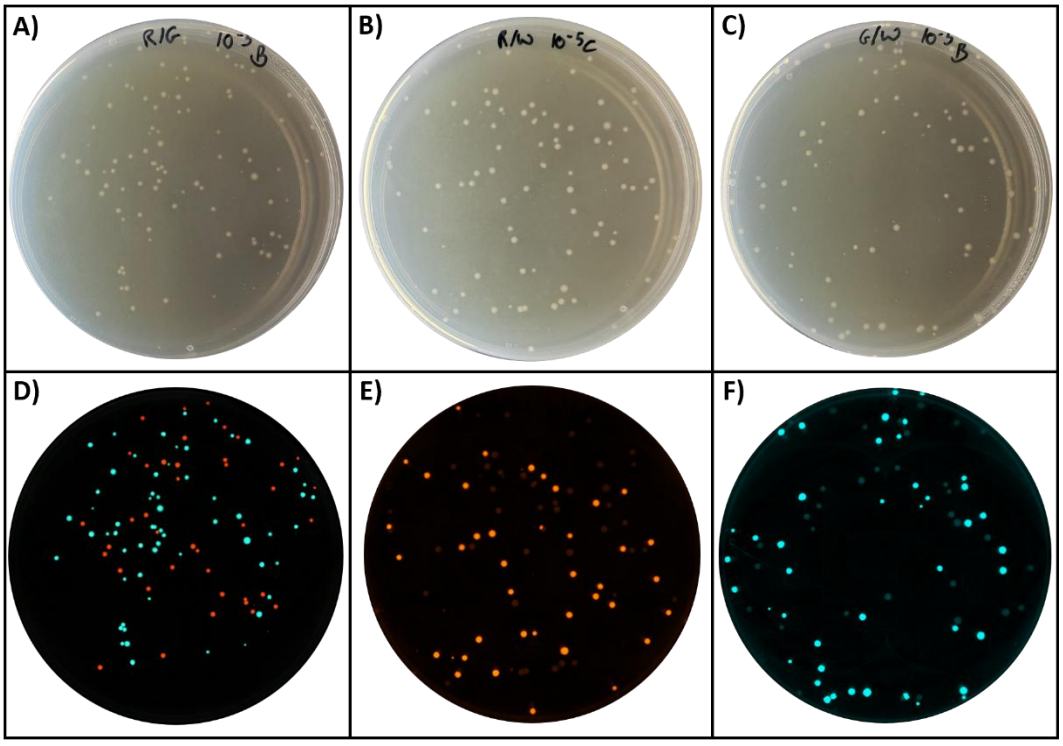


Figure 4.2: Appearance of mixed fluorescent and non-fluorescent *E. coli* strains.
Representative images of colony morphologies (A – C) and fluorescence (D – F) of different *E. coli* strains on nutrient agar plates. **A, D**) Mix of RFP and GFP expressing strains. **B), E)** Mix of RFP expressing and WT strains. **C), F)** Mix of GFP expressing and WT strains.

Plate counts of colonies showed no significant difference in CFU/mL between strains on RFP/GFP and GFP/WT plates (Figure 4.3, Panels A and C). For the RFP/WT plates mean CFU/mL was significantly higher for WT (5.25×10^7 +/- SE 0.375×10^7) compared with RFP

(3.84 x 10^7 +/- SE 0.189 x 10^7) cells (t-test = 3.337; d.f. = 4; p = 0.02891) (Figure 4.3, Panel B), which was unexpected given that the cultures had been mixed at equal ratios. Every colony on all of the GFP/RFP mix plates was fluorescent, indicating that the expression of both fluorescent marker genes was stable as this experiment was conducted without the addition of antibiotics.

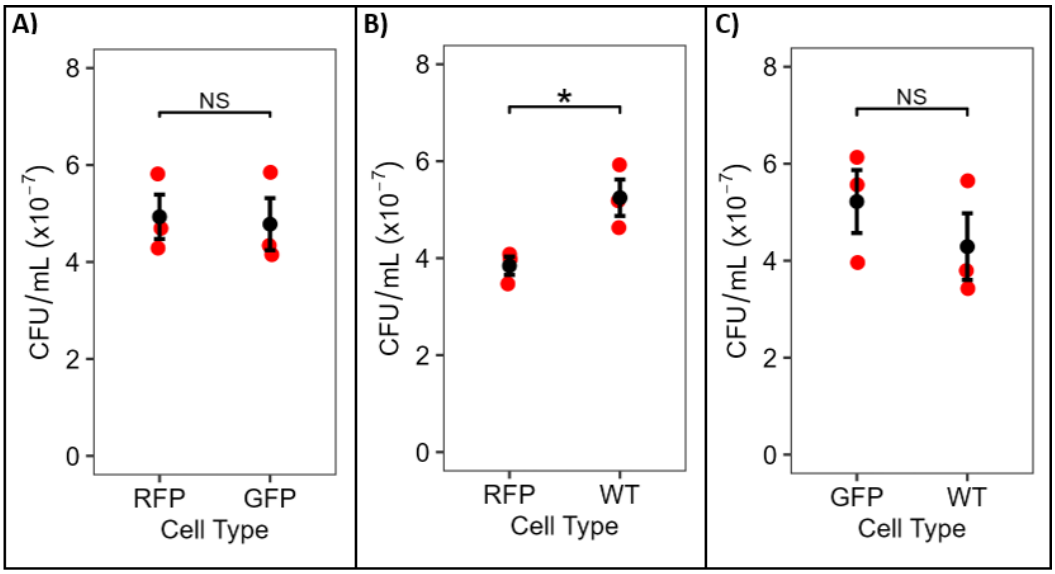


Figure 4.3: Pairwise *E. coli* strain mixing colony counts.
Colony forming units per mL of plated *E. coli* culture for the three different pairwise strain mixes. Means (black points) are derived from triplicate plate counts (red points). **A)** RFP and GFP strain mix. Mean colony counts were not significantly different between RFP (4.93 x 10^7 +/- 0.458 x 10^7) and GFP (4.78 x 10^7 +/- 0.39 x 10^7) cells; t = 0.215; d.f. = 4; p = 0.8399. **B)** RFP and WT strain mix. Mean colony counts were significantly higher in WT (5.25 x 10^7 +/- 0.375 x 10^7) compared with RFP (3.84 x 10^7 +/- 0.189 x 10^7) cells; t = 3.337; d.f. = 4; p = 0.02891. **C)** GFP and WT strain mix. Mean colony counts were not significantly different between GFP (5.22 x 10^7+/- 0.650 x 10^7) and WT (4.29 x 10^7 +/- 0.687 x 10^7) cells; t = 0.983; d.f. = 4; p = 0.3812. CFU values have been normalised to starting OD_{600} readings for each cell type. Error bars are +/- one standard error.

The results from the fluorescent strain mix plate counts (Figure 4.3, Panel A) indicated that there was no significant difference in the fitness between either strain in the conditions tested, suggesting that they might be suitable for use in direct competition experiments. To begin with experiments were conducted with the RFP strain.

4.2.1.2 Testing impact of resazurin/validation work

Before the RFP strain could be used in evolution experiments some additional validations of the assay system were required. Plating was used to generate an OD_{600} vs CFU curve to confirm a linear relationship between OD_{600} and cell counts below an OD_{600} of 0.75 (Figure A5, Appendix).

For the experimental evolution experiments it was essential that conditions remained anaerobic. This was because undetected loss of anaerobic conditions during evolution of strains could lead to the false conclusion that the strain had not successfully adapted to these conditions when it had not actually been exposed to them. To avoid this the coloured indicator resazurin would be used since it is very sensitive to oxygen status and changes colour from colourless to pink when oxygen is present (Guo *et al.* 2022).

To be a fair test only one parameter should be varied between test and control populations during directed evolution, which meant resazurin would need to be added to all culture vessels including controls containing oxygen, making them pink. This meant the test and control tubes used in this experiment would be different colours, which could present a problem given the method used to monitor growth (OD_{600} measurements) was based on spectrophotometry. An experiment was conducted to determine whether resazurin had a differential impact on OD_{600} measurements/bacterial growth in differing oxygen conditions.

The RFP strain was inoculated into growth medium with and without oxygen and resazurin. A set of uninoculated controls were included for each condition (Figure 4.4). As expected, the strain grew faster and to a higher final optical density in the presence of oxygen. Resazurin did increase OD_{600}, but this was relatively minor and (importantly) linear, demonstrated most clearly by the data for the uninoculated tubes with both resazurin and oxygen. Since no impact on growth had been observed and the increase in optical density measurements associated with resazurin in the presence of oxygen was linear it was deemed to be appropriate for use in further experiments.

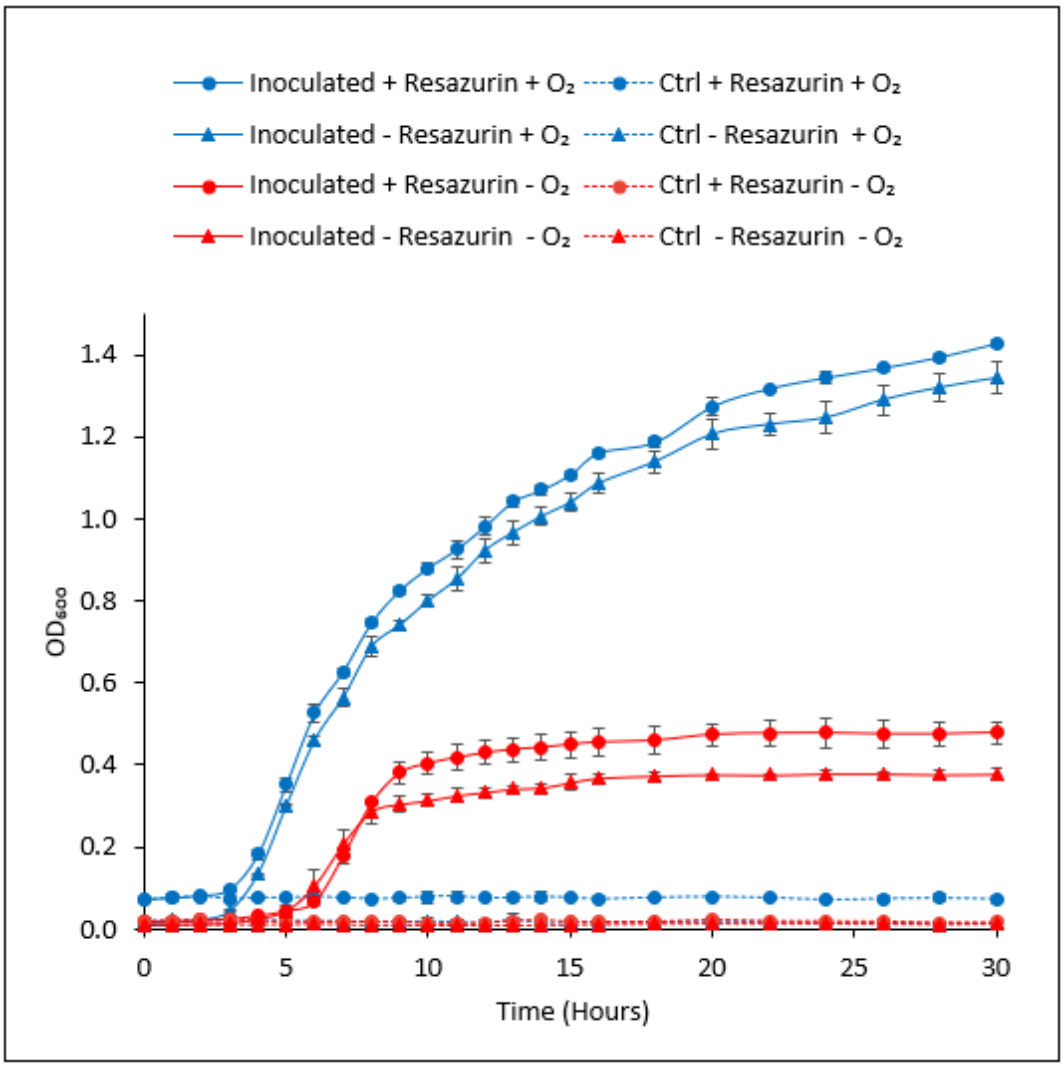


Figure 4.4: Effect of resazurin on *E. coli* culture turbidity readings.
OD_{600} measurements over 30 hours for *E. coli* RFP strain cultures (solid lines) and LB blank controls (dashed lines) with resazurin added (circles) or without (triangles) grown aerobically (blue) or anaerobically (red). Points are means of triplicate cultures. Error bars are +/- one standard deviation.

4.2.1.3 Evolving strains

The glycerol stock of the RFP strain (hereafter referred to as "ancestral strain") was used to generate eight replicate populations for directed evolution (Figure 4.5). Four were maintained in the test condition (no oxygen) and four in the control condition (with oxygen). All eight populations were passaged once a day into fresh medium at a 1:50 ratio. After one and two weeks an aliquot of each population was stored as a glycerol stock prior to transfer.

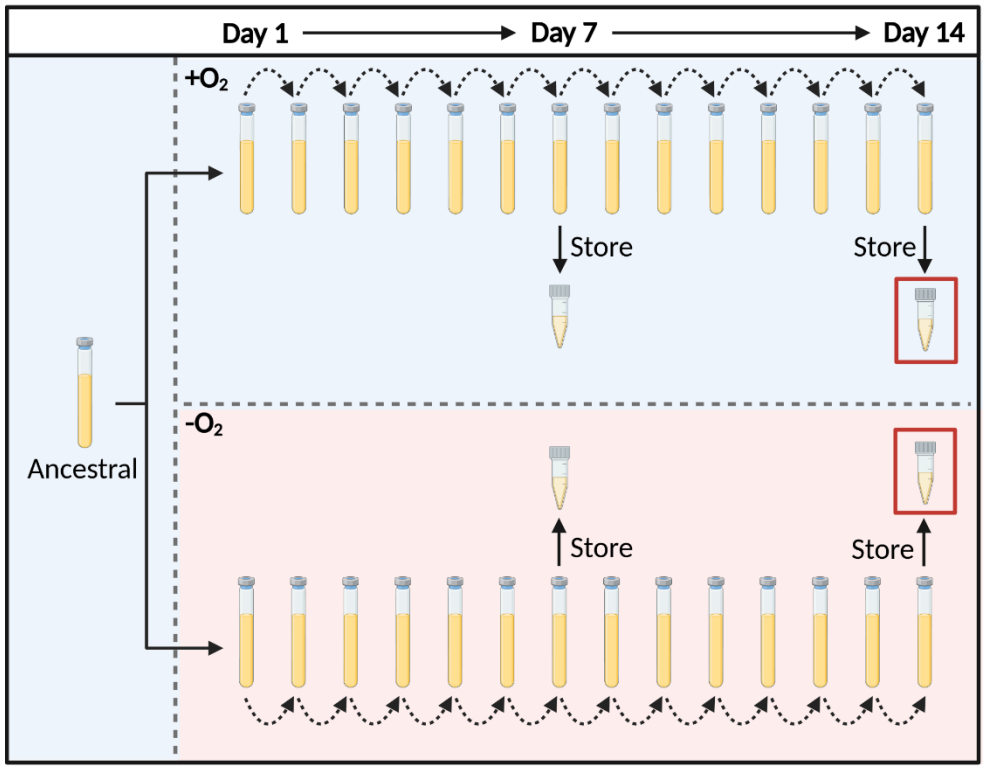


Figure 4.5: Directed evolution approach.
A culture of RFP expressing *E. coli* (Ancestral) is used to inoculate anaerobic or aerobic (control) medium to create distinct populations. Each population is passaged daily in its respective conditions (presence (blue box)/absence (pink box) of oxygen) for 14 days. Glycerol stocks of each population are taken on days seven and 14. The glycerol stocks taken on day 14 (red boxes) are used for the comparative growth assay (see below). Four separate aerobic and anaerobic populations were evolved: for clarity only one of each is shown.

Optical density of all eight replicate populations was recorded daily before transfer to detect any trends during the evolution process itself. For the anaerobic replicates (Figure 4.6, red lines) population 2 began to reach a higher final optical density than the other anaerobic cultures from day three, and this persisted until the end of the experiment. A similar increase was observed in population 1 on the final day of the experiment. Although there were no standout differences among aerobically passaged populations (Figure 4.6, blue lines), there was a general trend towards slight increases in final optical density as the experiment progressed.

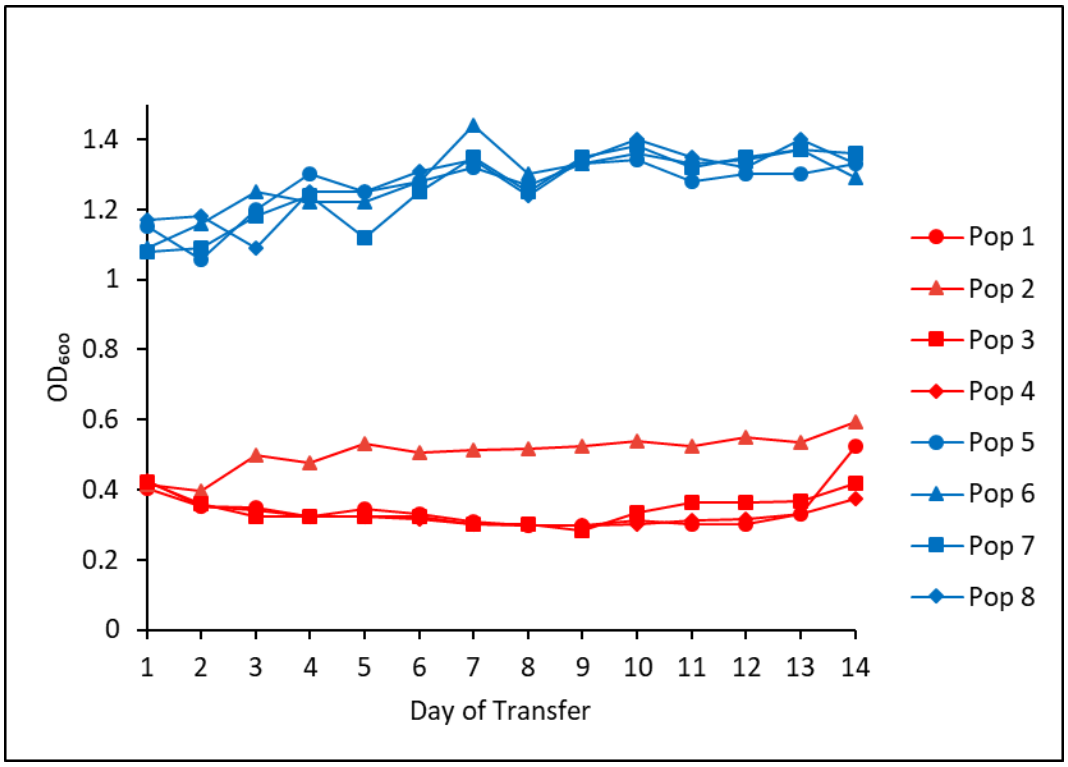


Figure 4.6: Directed evolution passaging.
OD_{600} measurements of the eight directed evolution populations of the *E. coli* RFP expressing strain taken each day for 14 days just before passage. Points are measurements of single cultures. Red points = anaerobic conditions, blue points = aerobic conditions.

4.2.1.4 Fitness assay

The glycerol stocks which had been generated after two weeks of directed evolution (Figure 4.6) were used to conduct a fitness assay, in which all eight strains were grown separately in the same conditions (anaerobically) and growth (determined by OD_{600} measurements) was used as a proxy for strain fitness. The ancestral and WT strains were

included as controls, as well as a mock inoculation. Prior to the experiment all eight evolved populations and the ancestral strain were streaked and inspected with fluorescent imaging, and every colony on every plate was confirmed to be red.

Anaerobically evolved populations performed better in general than those evolved aerobically in the fitness assay (Figure 4.7). Of note, anaerobic populations 1 and 2 grew more rapidly and to higher final densities than all other cultures, in agreement with their apparent improved performance during the evolution itself (Figure 4.6). Aerobic population 7 performed unexpectedly well, outperforming two of the anaerobically evolved populations. Anaerobic population 3 performed fourth best overall in the growth assay, while anaerobic population 4 and aerobic populations 5 6 and 8 all performed similarly, and only marginally better than the ancestral control strain (Figure 4.7). Neither control grew, which was expected as the mock inoculation contained no bacterial cells and WT is not resistant to chloramphenicol used in this experiment.

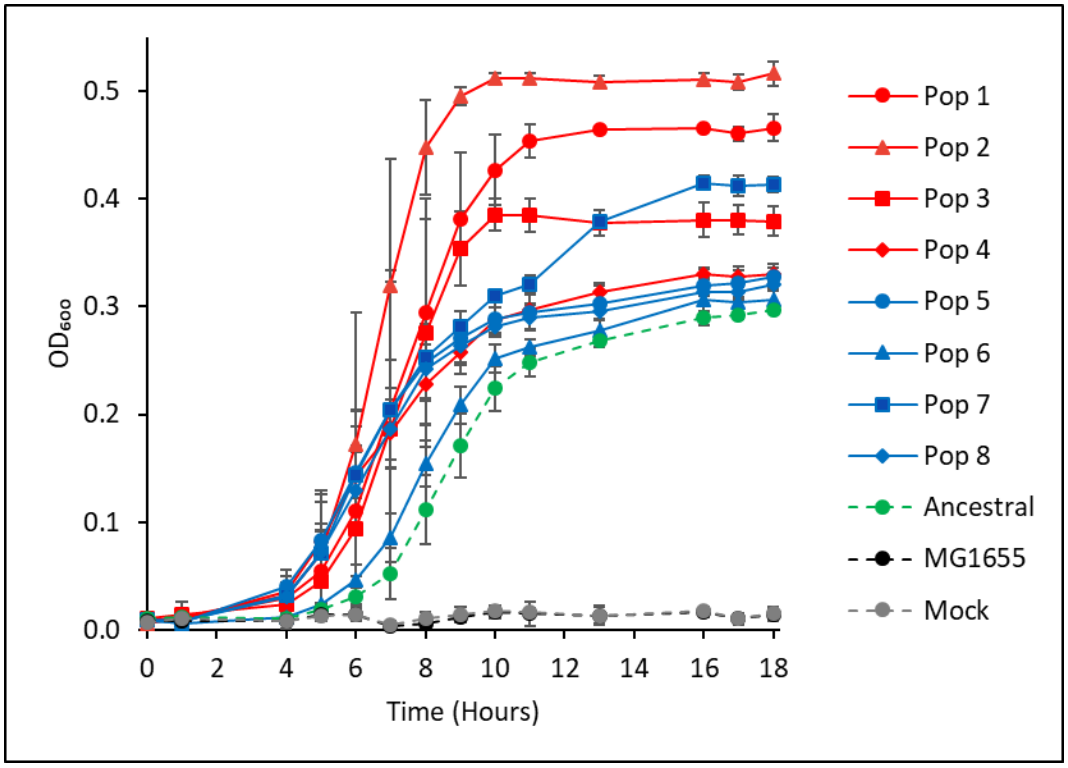


Figure 4.7: Comparative growth curves.
Anaerobic growth curves for the eight evolved *E. coli* RFP expressing populations. Populations 1-4 (red) had been evolved anaerobically. Populations 5-8 (blue) had been evolved aerobically. Controls (dashed lines) are the ancestral RFP expressing *E. coli* strain used to generate the eight evolved populations (green), WT *E. coli* MG1655 (black) and mock inoculated media (grey). Points are means of triplicate cultures. Error bars are +/- one standard deviation.

4.2.1.5 Strain sequencing

The outcome of the fitness assay had suggested that some populations appeared to be fitter than others in the test conditions. To investigate whether these differences had a genetic basis, DNA from all eight populations was submitted for whole genome sequencing followed by single nucleotide polymorphism (SNP) calling. To provide a point of reference the WT and ancestral strains were also included. Variant calling and filtering revealed a total of 12 mutated genomic loci across the evolved populations compared to that of the ancestor (Table 4.1). Average read coverage at these 12 positions was 1025 per population, indicating high confidence of the base calls at these positions.

Table 4.1: Mutant allele frequencies in experimentally evolved *E. coli* populations.
Mutations detected across the eight evolved *E. coli* populations compared to the ancestral MG1655_rLP5(nth-ydgR) strain from which they were derived (Anc) and a sequenced *E. coli* MG1655 isolate (WT). For mutations found in coding sequences the identity of the gene is provided. Mutations outside coding sequences are recorded as N/A. "Position" indicates the chromosomal coordinate of the mutation according to the NCBI *E. coli* K-12 MG1655 reference genome GCF_000005845.2. "Original" is the status of the allele in the ancestral strain (MG1655_rLP5(nth-ydgR)) and "Alternative" is the mutant version of the allele detected here. The frequency of each mutant allele in each population is given to the nearest 0.1%. Highlighted values are alleles found at 1% or higher frequency.

Mutation					Frequency by Population (%)									
Gene	**Position**	**Impact**	**Original**	**Alternative**	**Anc**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**WT**
dnaK	12604	Missense	G	A	0.0	0.0	0.0	0.2	0.0	0.0	0.0	**24.7**	**23.7**	0.0
dnaJ	14348	Stop	C	T	0.1	0.0	0.0	0.0	0.0	**20.7**	0.0	0.1	0.0	0.0
N/A	1978434	Unknown	AGGGGGGCT	AGGGGGGGCT	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**35.3**	0.0	0.0
flhC	1977596	Frameshift	TAAAAACT	TAAAACT	0.0	0.0	0.2	0.1	0.1	0.0	0.0	0.0	**13.8**	0.0
rcsC	2318555	Missense	C	T	0.1	0.0	0.1	0.0	0.1	0.1	0.0	0.0	**13.8**	0.1
crp	3486658	Missense	C	A	0.0	0.0	0.0	0.1	**21.5**	0.0	0.0	0.0	0.0	0.0
crp	3486664	Missense	A	G	0.1	0.0	0.1	0.0	**14.5**	0.0	0.0	0.0	0.0	0.0
frdA	4381253	Missense	A	G	0.0	**53.0**	**99.9**	0.3	0.0	0.0	0.0	0.4	0.1	**100**
treB	4465843	Missense	A	T	0.0	**3.3**	**54.3**	**87.6**	0.0	0.0	0.0	0.3	0.2	0.0
fimA	4543191	Missense	C	A	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**12.9**	0.0
fimA	4543373	Missense	A	C	0.0	0.1	0.0	0.0	0.1	**32.3**	0.0	0.0	0.7	0.0
fimA	4543604	Missense	G	A	0.0	0.0	0.0	0.0	0.1	0.0	0.0	**17.6**	0.2	0.0

In all but one case, mutant alleles were detected at low frequencies (predominantly less than 0.1%) in at least one population. The fact that this was also the case for both the WT and Ancestral populations suggests this was caused by some kind of sequencing assignment or alignment artefact since both DNA samples were derived from single streaked colonies so were genetically uniform. For this reason only mutations with a frequency of over 1% were considered further. Among the experimentally evolved populations 16 mutations were detected above 1% across the 12 genomic loci (highlighted values, Table 4.1).

No single mutation was present in the four populations that performed better in the growth assay (populations 1, 2, 3 and 7) and absent from those that did not (Figure 4.7) indicating that there was no single explanation for the observed differences. When considered collectively, some patterns in genetic variants present across the populations (Table 4.1) did appear to correspond to the growth data obtained in the fitness assay (Figure 4.7).

Only one mutation was present in both anaerobically evolved populations 1 and 2, which were the two best performers in the fitness assay (Figure 4.7). This A to G transition was near fixation in population 2, at a frequency of 53% in population 1 but absent from all other populations except for WT *E. coli* MG1655 (Table 4.1). The mutation results in a tyrosine to histidine change at position 356 in the fumarate reductase gene (*FrdA*) which is involved in anaerobic respiration in *E. coli* (Cecchini *et al.* 2002).

Three different mutations were detected in aerobically evolved population 7 (Table 4.1) which was the third best performing population in the fitness assay (Figure 4.7). 24.7% of this population had a missense mutation in the *dnaK* gene, which codes for a chaperone involved in protein folding in *E. coli* (Calloni *et al.* 2012). This variant was also present at a similar frequency in population 8, and a mutation was detected in *dnaJ* (which works as part of the same chaperone complex (Srinivasan *et al.* 2012)) in population 5 (Figure 4.7). Since neither population 5 nor 8 performed comparatively well in the growth assay (Figure 4.7) it seems unlikely the *dnaK* mutation was responsible for the increased growth phenotype observed in population 7.

Three distinct missense mutations were detected at different locations in the *fimA* gene in populations 5, 7 and 8 (Table 4.1), all of which had been evolved in aerobic conditions. *fimA* is part of the operon coding for type 1 fimbriae in *E. coli* (Klemm *et al.* 1985) which is involved in adhesion and biofilm formation (Isidro-Coxca *et al.* 2024). Again, since neither populations 5 or 8 performed as well as population 7 in the fitness assay it is unlikely the mutant status of this gene was responsible for the increased growth of population 7.

The final mutation detected in population 7 was a single base insertion near chromosomal position 1978434. The closest annotated gene to this location is *flhD,* which together with *flhC* (which is part of the same operon) acts as a master regulator for flagellar synthesis (Wang *et al.* 2006). It was found at a frequency of around 35% and was the only mutation detected outside of a protein coding sequence. As this variant was unique to population 7 it represents the most likely candidate for being responsible for the increased growth phenotype.

Anaerobically evolved populations 1, 2 and 3 had a missense mutation in the *treB* gene, which codes for a protein that transports trehalose into the cell (Klein *et al.* 1995). This variant was near fixation (87.6%) in population 3 (Table 4.1), which was the third best performing anaerobically evolved population (Figure 4.7). It was found at an intermediate level (54.3%) in population 2 (the best performer in the growth assay) and at a low level (3.3%) in population 1 (the second best performer).

Population 4 was the worst performing anaerobically evolved population (Figure 4.7). Two variants were detected at relatively low levels at two different points in the *crp* gene. The gene product is the cyclic-AMP receptor protein, a transcription factor that acts as a master regulator for many other genes in *E. coli* (Green *et al.* 2014).

Variants in two genes were detected only in population 8 (Table 4.1). An insertion inside the coding sequence of the *flhC* gene, which is involved in flagellar regulation (Wang *et al.* 2006) as discussed above, was detected at a frequency of just under 14%. This mutation results in a frameshift from leucine 83 in the 192 amino acid protein (Wang *et al.* 2006), impacting the majority of its length. A missense mutation was detected at around 14%

frequency in *rcsC*, the gene product of which is part of a signalling pathway involved in sensing membrane integrity, and has a number of outputs including cell surface remodelling (Clarke 2010; Meng *et al.* 2021). None of the variants detected in population 8 appeared to exert a substantial impact on growth since this population performed only marginally better than the ancestral control population (Figure 4.7) as was the case for all other aerobically evolved populations except population 7.

4.2.2 Obtaining chassis organisms: plating

The other approach selected for obtaining more robust chassis organisms was to attempt to obtain one or more close relatives of laboratory strains from a wild AD community.

4.2.2.1 Selective plating from sludge

4.2.2.1.1 *In silico* identification of candidate organisms

Existing datasets from previous projects in the lab were first used to try to identify candidate organisms to isolate which had previously been found to be active (i.e. synthesising proteins) and present in a broad range of digesters (see section 2.3.2, Methods). An initial "top down" approach in which whole genome datasets were used to generate a list of organisms present across the conditions of interest did not provide enough taxonomic resolution to identify candidate species. An alternative "bottom up" approach was used to create a search list of 22 desirable candidate organisms which had existing genetic tools, and were also likely to be present in AD (Table A2, Appendix). Searching for these organisms in whole genome data presented similar problems to those explained above, but when the approach was adapted to instead be based on 16S sequences one of the "desirable" organisms, *Clostridium beijerinckii* (*C. beijerinckii*) was putatively identified in the datasets. As these approaches had not yielded many candidate organisms to isolate, two additional organisms, *E. coli* and *B. subtilis*, were added to the list as they are well developed facultatively anaerobic model organisms (Nakano and Zuber 1998; Conway and Cohen 2015; Errington and Aart 2020) that could be present in AD.

4.2.2.1.2 Attempted isolation of chosen organisms from sludge

Selective plating protocols and suitable media for isolation of the three chosen candidate organisms were obtained. Membrane lactose glucuronide agar (MLGA) is selective for coliforms: *E. coli* should appear green on plates composed of this medium, while non-target organisms are pink (The Standing Committee of Analysts 2016). Nutrient agar supplemented with bromothymol blue (NAB) selects for aerobic spore forming bacteria including *Bacillus* species, which should appear as pale green colonies on these plates (Environment Agency 2007). Tryptose sulphite cycloserine agar (TSC) is selective for sulphite-reducing bacteria which typically appear as black colonies on plates and has been reported to isolate *C. beijerinckii* (The Standing Committee of Analysts 2021). A fresh

sludge sample was plated on these three different selective plates as well as Schaedler agar (SA), a non-selective medium included as a control (Figure 4.8).

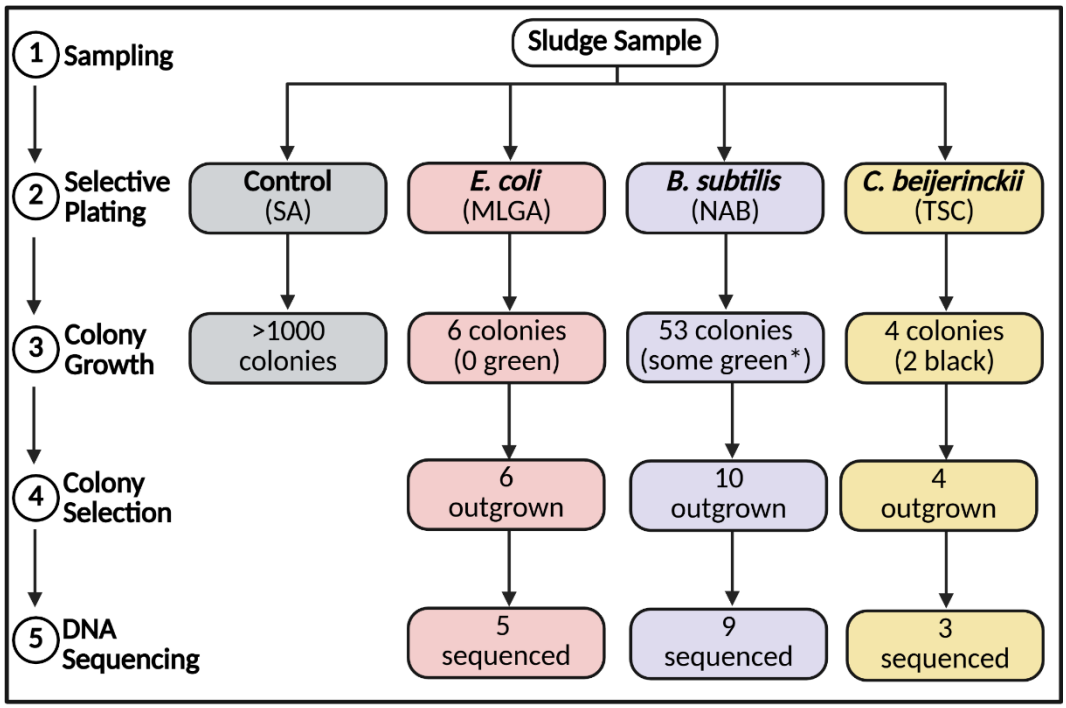


Figure 4.8: Selective plating overview.
1) A sample of sludge from the Esholt site is taken. **2)** The sample is plated on three types of selective media, each intended to enrich one of the desired organisms. A control (non-selective) plate type is also included to ensure viability of the sample. **3)** Plates are incubated and the number of colonies obtained across all plates of each type is counted (or estimated in the case of the control). The number of colonies which were the expected colour for the desired organisms on selective plates are indicated in brackets. **4)** Colonies are chosen from the selective plates and outgrown in liquid culture, then DNA is isolated. **5)** The majority of DNA samples are extracted for sequencing for identification. SA = Schaedler agar. MLGA = membrane lactose glucuronide agar. NAB = nutrient agar with bromothymol blue. TSC = tryptose sulphite cycloserine agar. * *Bacillus* colonies should appear light green on NAB plates, however due to very subtle colour differences it was not possible to quantify how many of the 53 colonies on these plates were this colour.

Colonies on MLGA and NAB plates had to be incubated for an additional day as colonies were small and poorly grown after 24 hours (the maximum suggested incubation period for both protocols). Plate counts performed at this point revealed there were substantially fewer colonies on the selective plates compared with the non-selective control (Figure 4.9) as would be expected. Among the selective media most colonies were found on NAB (53 across all plates) followed by MLGA (six colonies) then TSC (4 colonies).

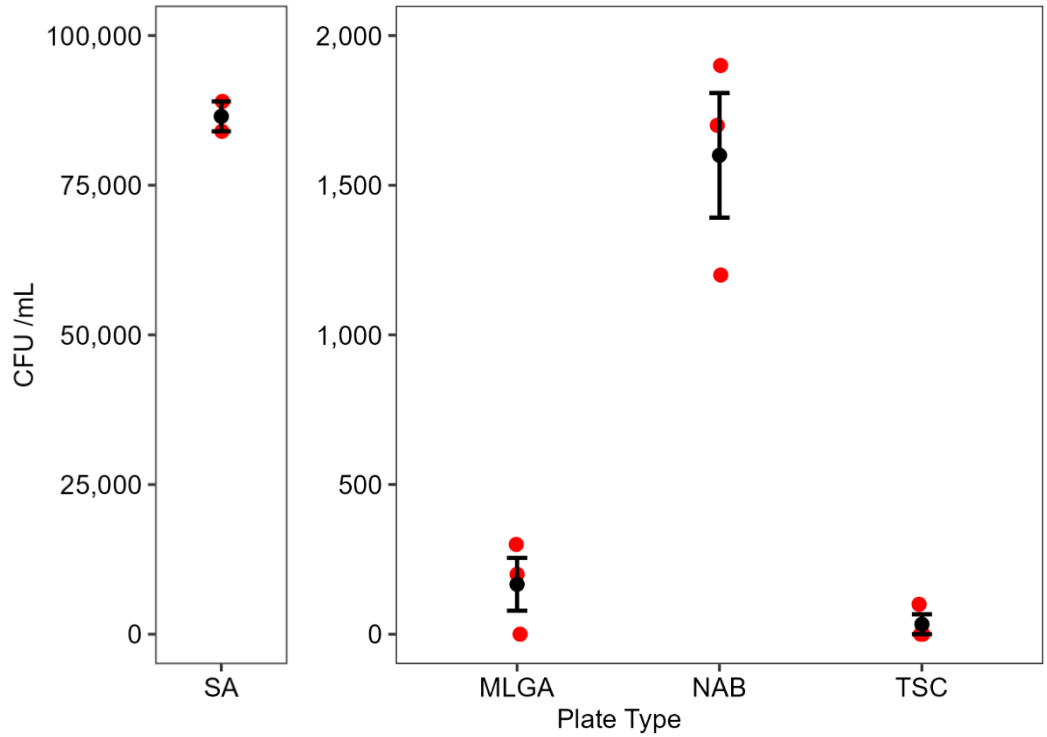


Figure 4.9: Colony counts of selective vs non-selective plates.
Mean colony forming units per mL of wet sludge (black points) calculated from triplicate plate counts (red points) of non-selective (SA) vs selective (MLGA, NAB, TSC) plates. Error bars are +/- one standard error. SA = Schaedler agar, MLGA = membrane lactose glucuronide agar, NAB = nutrient agar with bromothymol blue, TSC = tryptose sulphite cycloserine agar.

All six of the colonies recovered on MLGA plates were pink, indicating they were non-target organisms. Since these plates had been incubated for longer than suggested in the protocol it was possible the colour indication had not worked properly, so all six were inoculated into liquid culture for DNA extraction and subsequent identification through sequencing.

All four colonies from the TSC plates were outgrown in liquid culture for DNA extraction. Two of the four colonies appeared black on plates which would be expected for *Clostridia*. Neither of these colonies grew further in the presence of oxygen when removed from the anaerobic chamber, further indicating they might be *Clostridia* as they are obligate anaerobes (Morvan *et al.* 2021).

For NAB plates *Bacillus* colonies should be light green, but this was found to be extremely challenging to gauge by eye, so 10 colonies encompassing a variety of morphologies were

selected for outgrowth. Six morphologically identical colonies were also present on one of the control NAB plates. One of these was selected for outgrowth and sequencing to determine the identity of this contamination.

Since the NAB and MLGA plates had not performed as expected, one plate of each type was spread with a positive and negative control organism. Both organisms grew as expected and were the correct colours on each plate type (Figure 4.10) indicating that there was no problem with the plates themselves.

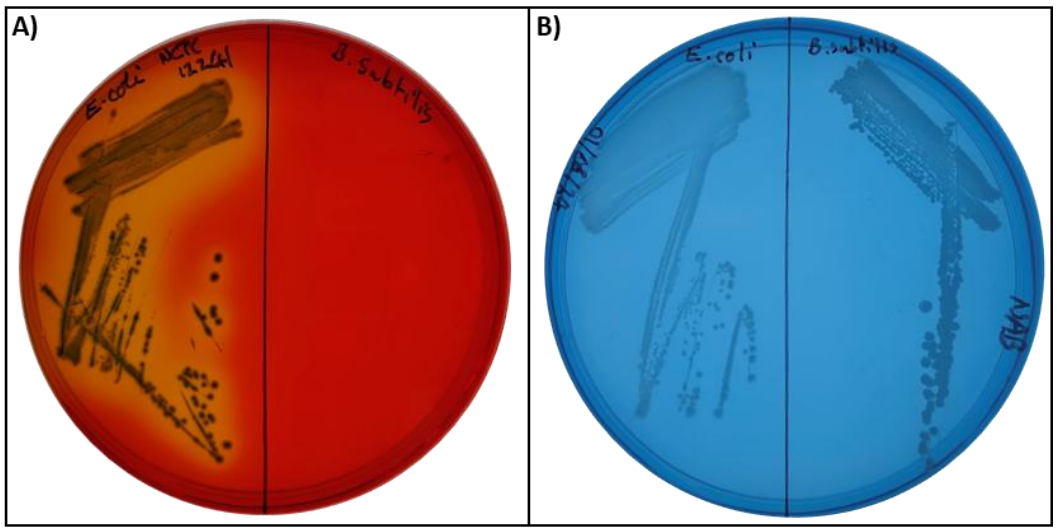


Figure 4.10: Selective plate controls.
Control plates for **A)** MLGA and **B)** NAB. In both cases *E. coli* is streaked on the left sector and *B. subtilis* on the right.

DNA was extracted from all liquid cultures from the selected colonies using an appropriate DNA extraction protocol (i.e. including a pre-heating step for presumptive *Bacillus* and *Clostridial* cultures from NAB and TSC plates as these organisms can form spores). One culture from each media type was not taken forward for sequencing due to slow growth or insufficient DNA for sequence submission. The remainder of the cultures (five derived from MLGA plates, nine from NAB and three from TSC) were submitted for full genome sequencing to enable identification.

The short read data derived from each sample was assigned taxonomic identity using Kraken2. For some samples the majority of the reads aligned to a single organism down to the species level (e.g. sample one from NAB plates: Figure 4.11, Panel A). Others did

not provide species level resolution. For example, in sample three from NAB plates the majority of the reads are assigned to *Burkholderiaceae* down to the family level (Figure 4.11, Panel B) but there was substantial branching at higher taxonomic resolution, making identification beyond the family level challenging.

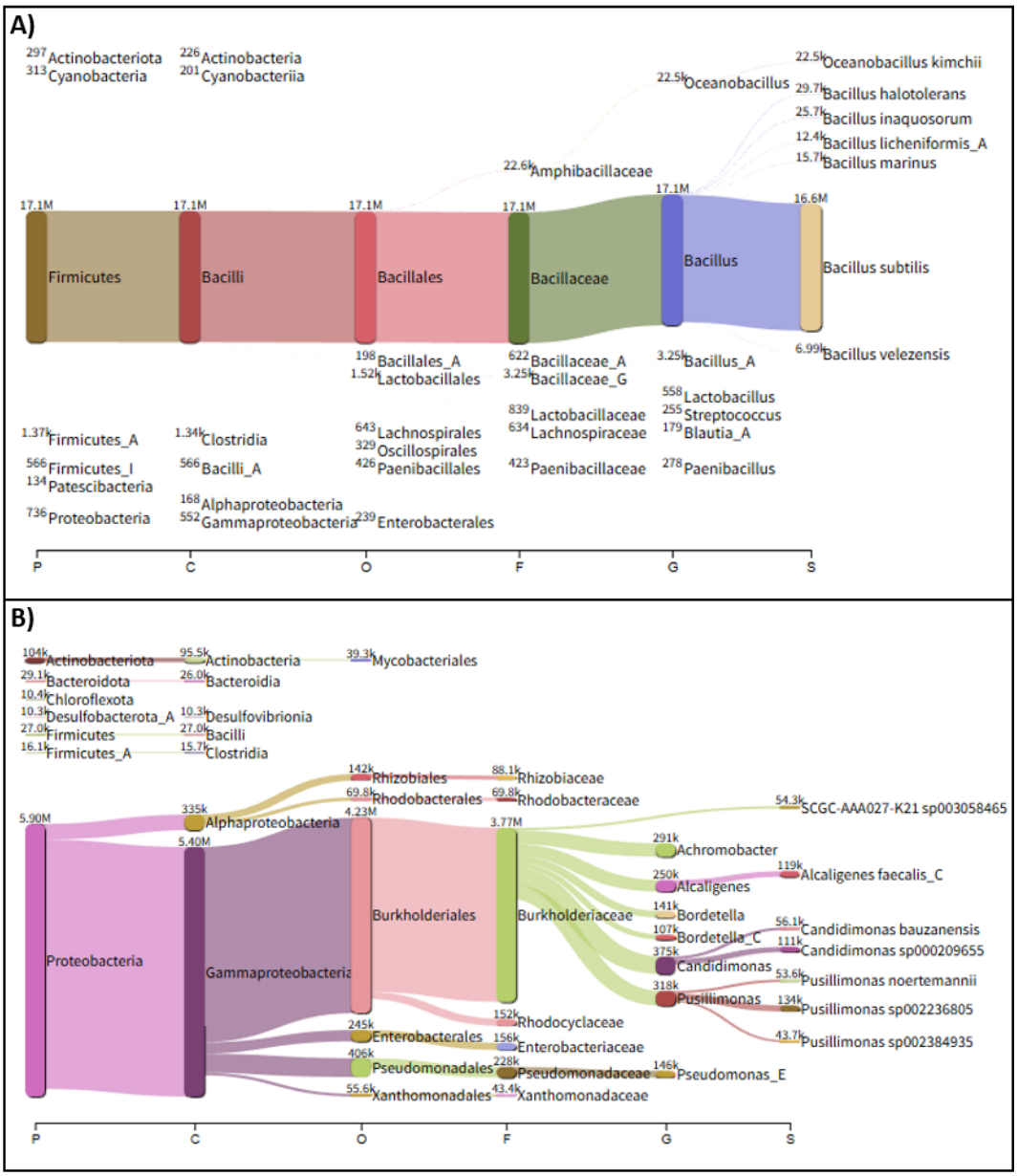


Figure 4.11: Taxonomic classification of short read selective plate DNA samples. Representative Sankey plots for Kraken2 taxonomic assignment of whole genome sequencing reads from DNA derived from **A)** NAB 1 and **B)** NAB 3. Bars indicate taxonomic assignment of reads from the phylum (left hand side) to the species (right hand side) level (see scale bar at bottom). The number of reads assigned to each classification is indicated above the respective node, with the thickness of the bars giving an indication of the number of reads. NAB = nutrient agar with bromothymol blue. Scale bar: P = phylum, C = class, O = order, F = family, G = genus, S = species. Diagrams produced in Pavian.

10 samples were identified in this manner to species level, including all five samples from MLGA plates, two of the three samples from TSC plates and two of the nine from NAB (Table 4.2). The contaminant colony on the NAB control plate was identified as

Acinetobacter scindleri, but apart from a single sample (NAB 4) the other organisms on NAB plates were distinct from this. *B. subtilis,* which was one of the target organisms, was isolated on the appropriate plate type in one case (NAB sample 1).

Table 4.2: Identification of selective plating DNA samples.
Inferred identity of distinct organisms based on DNA sequencing of selective plate samples. "Plate type" is the media from which the respective sample was originally isolated. Identity is based on analysis of Kraken2 taxonomic assignments of unassembled whole genome sequencing reads. Brackets after predicted identities indicates taxonomic rank: S = species, G = genus, F = family. MLGA = membrane lactose glucuronide agar, NAB = nutrient agar with bromothymol blue, TSC = tryptose sulphite cycloserine agar.

Plate Type	Sample	Identity
MLGA	1	*Pseudomonas guangdongensis* (S)
MLGA	2	*Pseudomonas guangdongensis* (S)
MLGA	3	*Pseudomonas guangdongensis* (S)
MLGA	4	*Alcaligenes faecalis* (S)
MLGA	5	*Pseudomonas guangdongensis* (S)
NAB	1	*Bacillus subtilis* (S)
NAB	2	*Burkholderiaceae* (F)
NAB	3	*Burkholderiaceae* (F)
NAB	4	*Burkholderiaceae* (F)
NAB	5	*Acinetobacter* (G)
NAB	6	*Stenotrophomonas* (G)
NAB	7	*Burkholderiaceae* (F)
NAB	8	*Burkholderiaceae* (F
NAB	9	*Bacillus lichenformis* (S)
NAB Ctrl	1	*Acinetobacter scindleri* (S)
TSC	1	*Paraclostridium masiliensis* (S)
TSC	2	*Paraclostridium masiliensis* (S)
TSC	3	*Pseudescherichia* (G)

4.2.2.2 Untargeted isolation of organisms from sludge

An alternative approach to finding possible AD chassis organisms is simply to plate sludge non-selectively and identify isolated organisms.

4.2.2.2.1 Simple sludge plating conditions

A chassis organism should be easy to grow (i.e. not requiring complexed medium) so the isolation effort was begun with nutrient agar/broth, a widely used and simple medium. The intention had been to plate and incubate sludge both aerobically and anaerobically, but due to a chamber malfunction only aerobic work could be done for this experiment. Fresh Esholt sludge was plated and colony counts and morphologies were recorded daily for five days (Figure 4.12). Colony counts increased each day, and existing colonies grew in size. Nine morphologically distinct colonies picked on day one were outgrown in liquid culture and stored as glycerol stocks. Morphological distinctness of existing colonies was found to increase over time, so 18 further colonies were picked and processed for storage on day four.

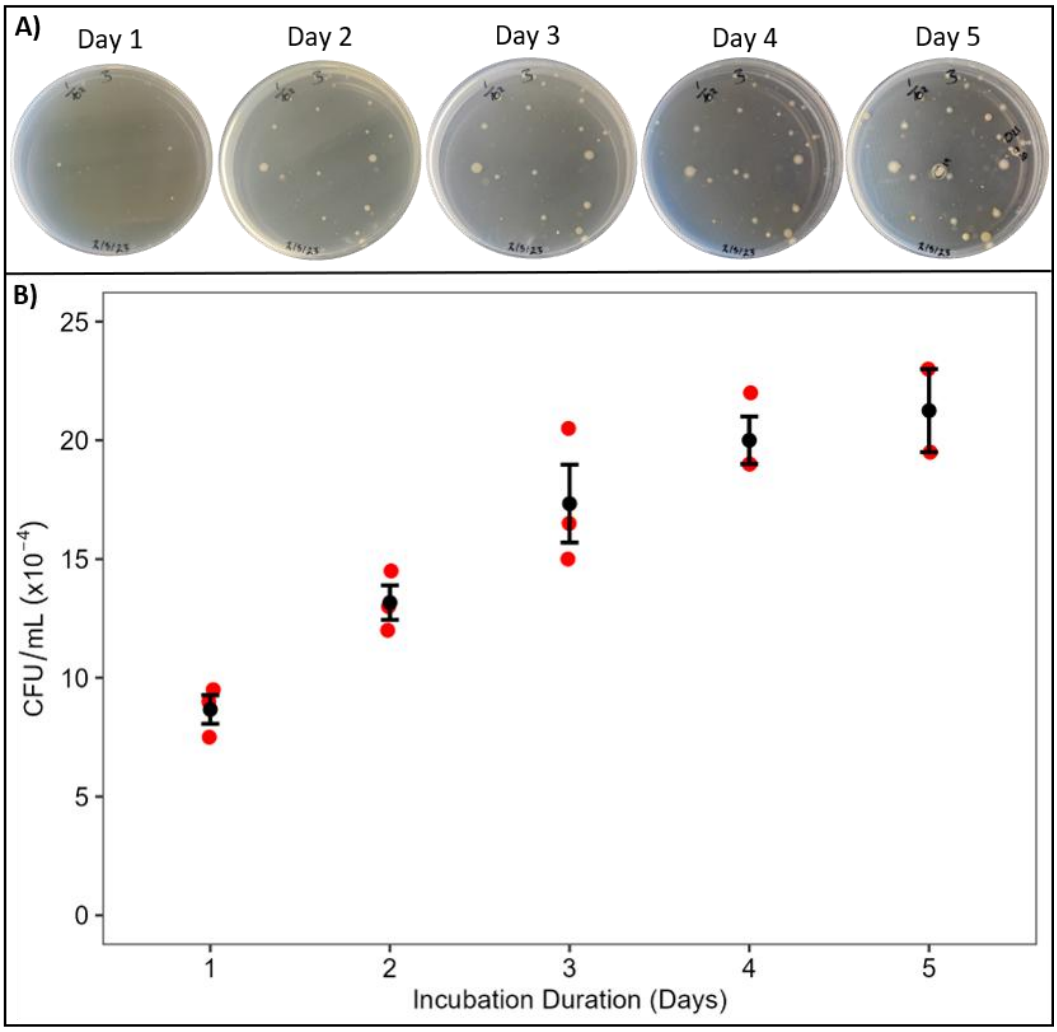


Figure 4.12: Plated sludge colony morphologies and counts over time.
A) Representative images of changes in colony morphologies on a single nutrient agar plate over five days. **B)** Mean colony forming units per mL of wet sludge (black points) calculated from triplicate plate counts (red points) taken over five days. Error bars are +/- one standard error.

This preliminary work revealed a number of useful insights which streamlined subsequent isolation experiments. Firstly, new colonies arose each day so longer incubation periods mean increased numbers of (presumably slower growing) organisms. Colonies also became easier to visually distinguish as incubation time progressed, with a marked increase in morphological distinctness from day two. Beyond four days of incubation some colonies had grown to an extent where they began to merge, and the medium also appeared visibly drier by this point, so a four day incubation was selected as a compromise between these factors. For storage of picked colonies, length of time of incubation in liquid culture prior to glycerol storage for one or four days did not make a difference in

recovery success rate as assessed by revival of stocks on streak plates. This test also revealed that in some cases organisms could be recovered even if the liquid culture used to make that stock had not had visible signs of growth (i.e. no apparent growth in liquid culture did not necessarily mean absence of organisms).

4.2.2.2.2 Broadening sludge plating conditions

Experiments were carried out to assess whether including a broader range of media types as well as anaerobic vs aerobic incubation had an impact on number and/or visual diversity of microorganisms obtained. Fresh sludge was plated on three different media types: nutrient agar (NA), brain heart infusion agar (BHI) and reinforced clostridial agar (RCA), each under anaerobic and aerobic conditions. All plates were incubated in their respective conditions for four days, and colonies were counted. A subset of those judged to be morphologically distinct (33 aerobic, 34 anaerobic) were isolated via outgrowth in their corresponding broth and stored as glycerol stocks with or without oxygen. For each media type and condition, a glycerol stock was prepared to control for sterility.

Colonies on aerobically incubated plates were visually distinct and easy to count (Figure 4.13, Panel A). Significantly more colonies were recovered on NA compared with BHI and RCA plates (Figure 4.13, Panel B). 33 morphologically distinct colonies were picked and outgrown in their respective liquid medium. 32 had visibly grown after four days of liquid culture, and all were stored as glycerol stocks for later revival and identification.

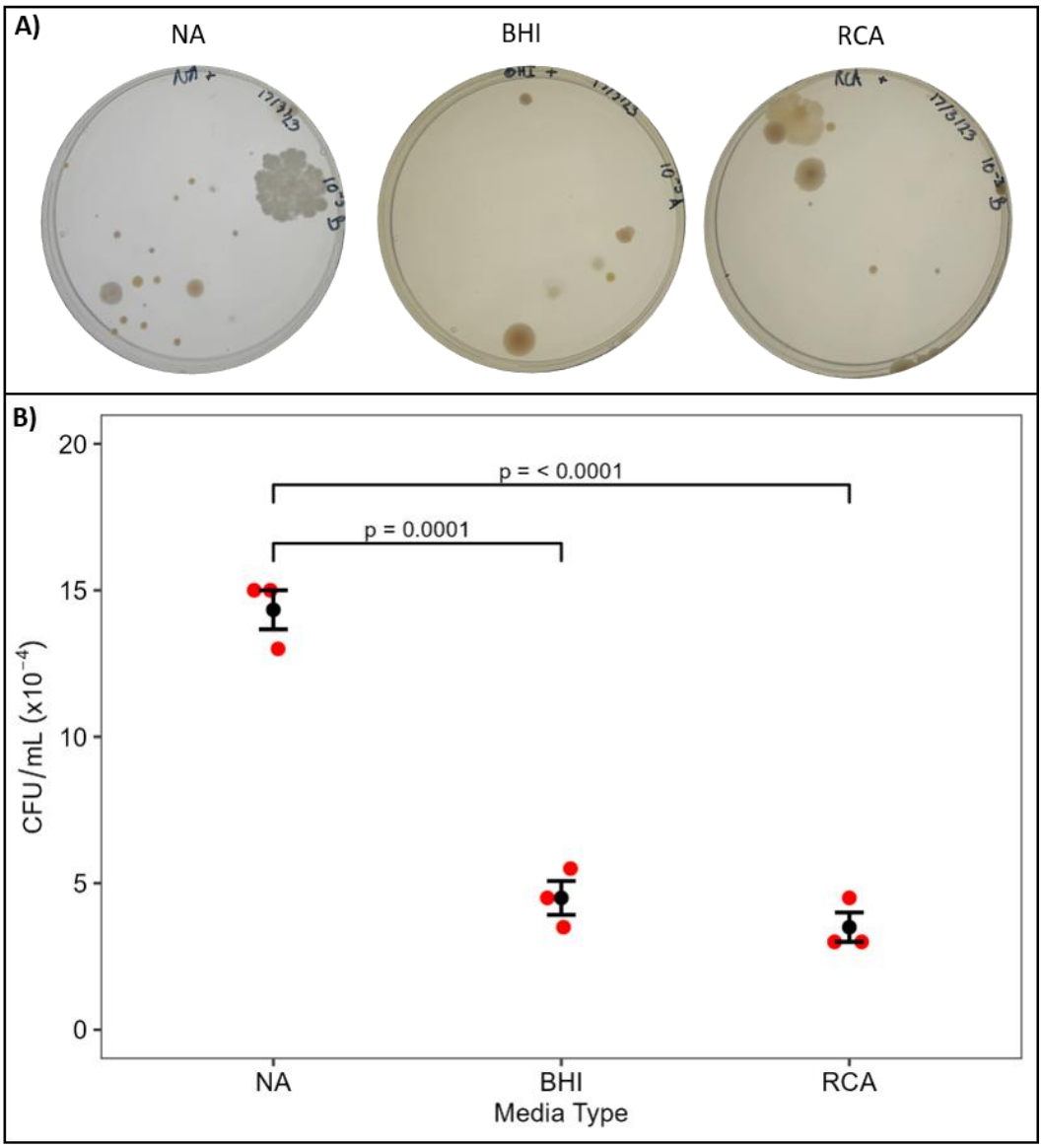


Figure 4.13: Aerobic plated sludge colony morphologies and counts between media. **A)** Representative images of colonies on the three different media used. **B)** Mean colony forming units per mL of wet sludge (black points) calculated from triplicate plate counts (red points) after four days of aerobic incubation. There was a significant difference in number of colonies between media types ($F = 104.6$; d.f. = 2, 6; $p = 0.00002$). Post-hoc testing showed significantly higher CFU for NA (mean 14.3×10^4 CFU +/- SE 0.667×10^4) compared with both BHI (4.5×10^4 +/- 0.577×10^4 CFU; $p = 0.0001$) and RCA (3.5×10^4 +/- 0.5×10^4 CFU; $p = < 0.0001$). Error bars are +/- one standard error. NA = nutrient agar. BHI = brain heart infusion agar. RCA = reinforced clostridial agar.

Colonies on anaerobically incubated plates (Figure 4.14, Panel A) were generally smaller than those on aerobic plates, and all plates including controls also contained large numbers of small translucent colonies. Only those visible under a light box were counted (Figure 4.14, Panel B). Due to the presence of translucent colonies and the comparatively small size of those visible under a light box, these counts should be viewed with caution. Comparison between anaerobically incubated media types can be made as all were subject to the same difficulties: no significant differences were found between media types (ANOVA: $F = 2.111$; d.f. = 2, 6; $p = 0.2023$) (Figure 4.14, Panel B). Comparisons between counts on aerobic (Figure 4.13, Panel B) and anaerobic plates (Figure 4.14, Panel B) cannot be reliably made as there were many more colonies present on the anaerobic plates, these were just not included in the counts.

The presence of contamination on control plates could be explained by the fact that it was not possible to work under completely sterile conditions in the anaerobic chamber. Conversely, aerobic plate manipulations could be performed in a laminar flow hood which would explain why no contamination was detected on these plates. 34 of the larger visibly distinct colonies were picked and outgrown anaerobically in the corresponding liquid medium. 20 visibly grew, but all 34 cultures were stored as glycerol stocks.

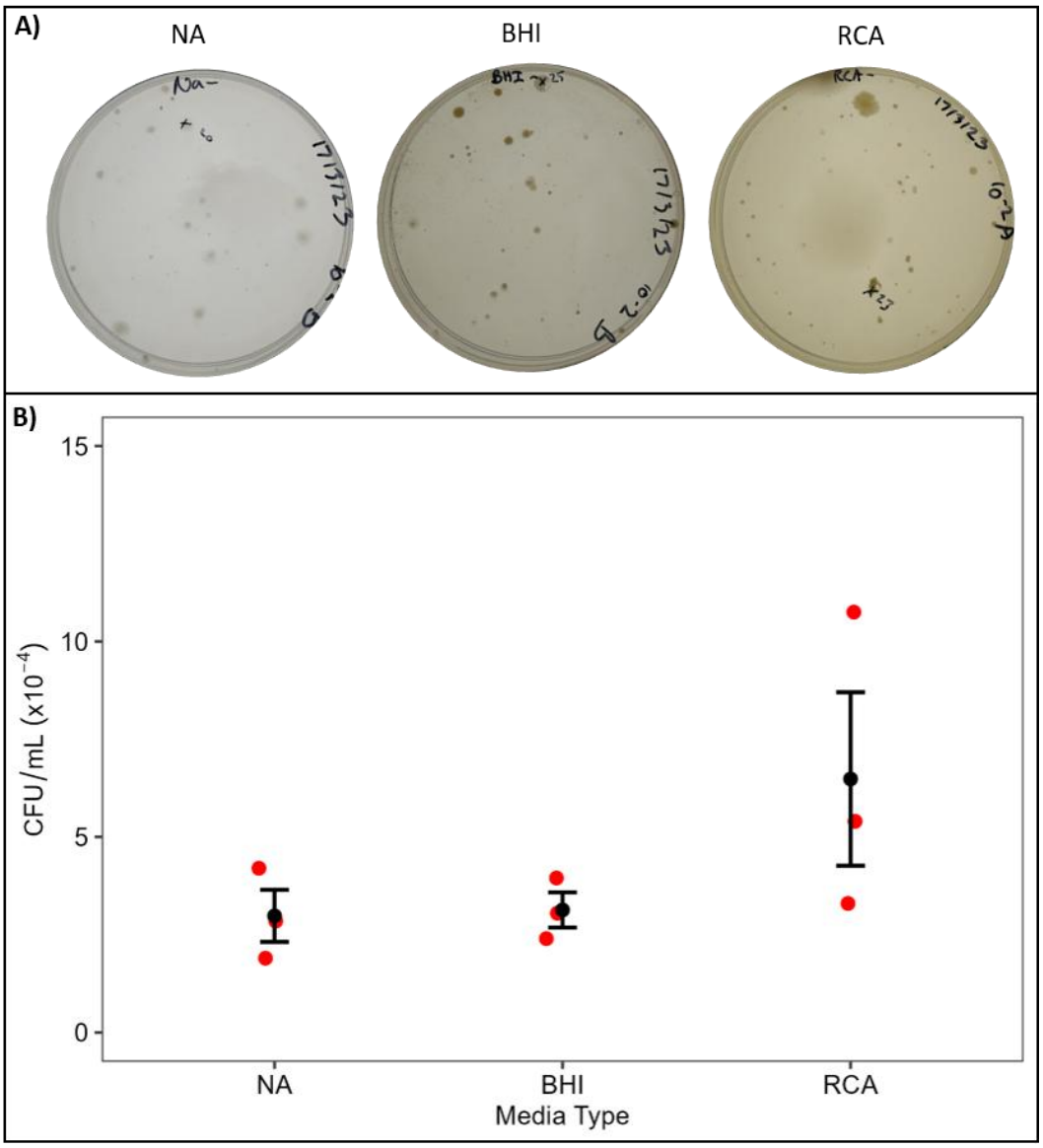


Figure 4.14: Anaerobic plated sludge colony morphologies and counts between media. **A)** Representative images of colonies on the three different media used. **B)** Mean colony forming units per mL of wet sludge (black points) calculated from triplicate plate counts (red points) after four days of anaerobic incubation. There was no significant difference in number of colonies between media types (F = 2.111; d.f. = 2, 6; p = 0.2023). Error bars are +/- one standard error. NA = nutrient agar. BHI = brain heart infusion agar. RCA = reinforced clostridial agar.

4.2.2.2.3 Recovery of unidentified isolated organisms

Around 100 different colonies had been preserved as glycerol stocks. To determine how well these organisms survived storage revivals were carried out by restreaking on solid media plates in the conditions in which they had been isolated.

33 aerobic and 34 anaerobic colonies isolated on the same day and stored at -80 °C for one month were restreaked in their respective conditions. Of these, 32 aerobic and 25 anaerobically stored organisms were successfully recovered on plates after four days of incubation. Control anaerobic glycerol stocks also yielded colonies on plates, likely due to the contamination issues encountered during plating. The fact that no colonies grew on some of the attempted anaerobic colony revivals indicated that the origin of this contamination was likely from the control glycerol stocks themselves rather than contamination introduced during this restreaking.

A further set of revivals of aerobic stocks was performed after 13 months in cryostorage. While more stocks were revived at this later timepoint, this included the same 33 stocks which had also been revived after one month. On both occasions all but the same single organism was successfully recovered by plate streaking, indicating that the additional 12 months in storage had not reduced viability for any of the original cultures as determined by this method.

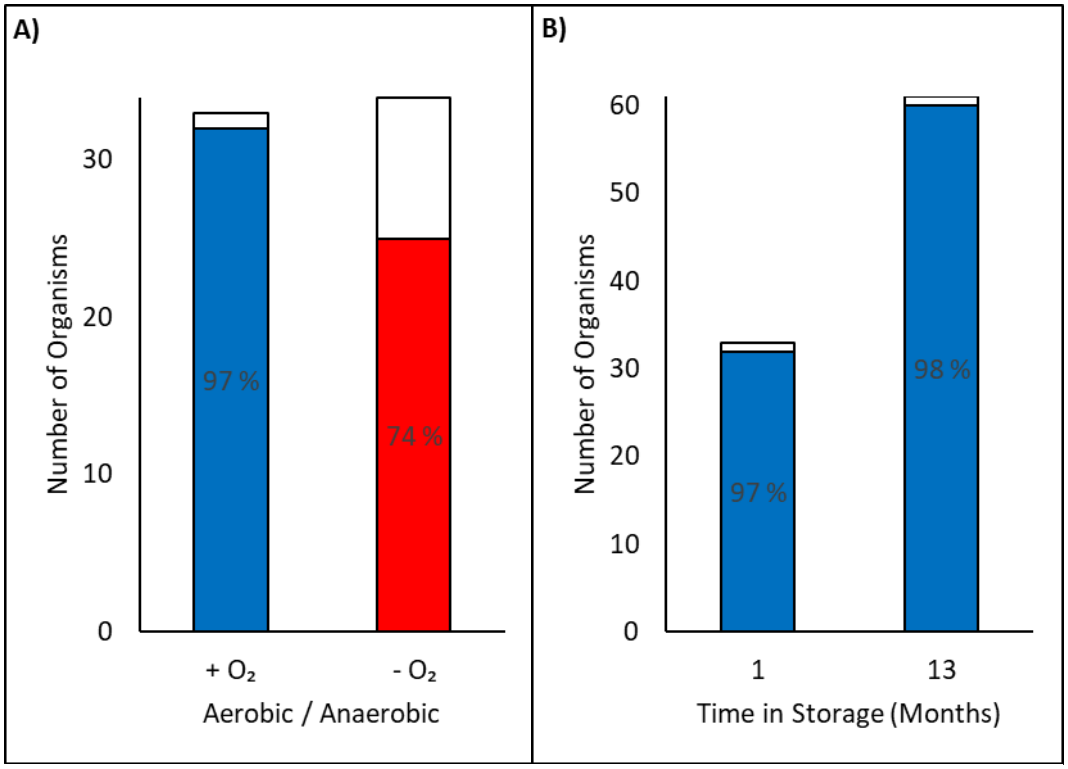


Figure 4.15: Microbial recovery success rates following cryostorage.
Revival success rates determined by plate streaking for unidentified microbial cultures stored as glycerol stocks at -80 °C. Entire bars indicate the total number of glycerol stocks for each category for which revival via plate streaking was attempted for the X axis condition. Coloured portions of the bars indicate the number of successful revivals in that category. Bar annotations are percentage of attempted revivals that were successful. **A)** Success rates for revivals after approximately one month in cryostorage for unidentified organisms isolated and stored either aerobically or anaerobically. **B)** Success rates for revivals of aerobically isolated and stored organisms after one and 13 months in cryostorage. Bar heights are different because only a subset of the aerobically stored and isolated organisms were revived after one month in storage.

4.2.3 Identifying isolated organisms

A method of identification based on DNA sequencing was devised for the stored organisms (Figure 4.16). The method was designed based on the dimensions of a 96 well plate to enable culture quantities to be easily scaled down if it was found to be successful. The approach was based on the principle of inoculating and growing 96 unidentified organisms in liquid culture separately, pooling and sequencing rows and columns, and using overlaps in the sequencing data to back calculate the contents of each intersection (Figure 4.16). If successful, this would reduce the number of individual samples needed for extraction and sequencing from 96 to 20, reducing labour and costs approximately fivefold.

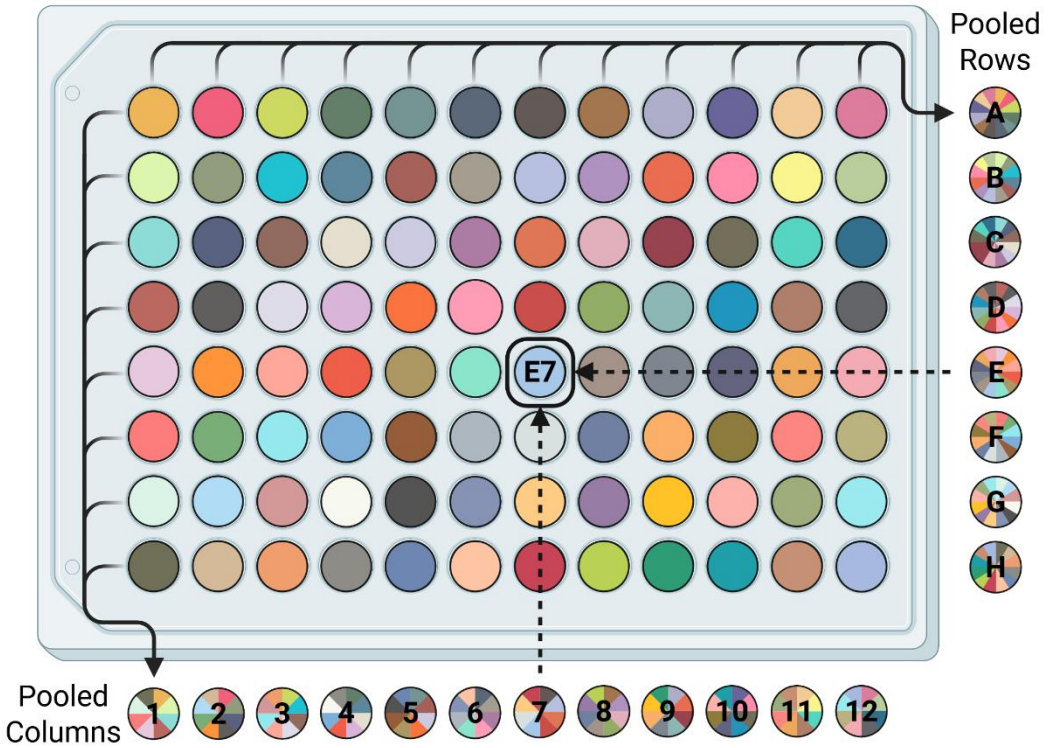


Figure 4.16: Proposed organism identification strategy.
96 unidentified microorganisms are grown in separate wells in a 96-well plate (coloured circles). Following growth, organisms are pooled by columns and rows: 12 column pools each containing a mix of eight organisms and eight row pools each containing 12 organisms (20 pools total). DNA is extracted from pools and sequenced. Sequence alignments of assemblies from rows and columns is used to identify the organisms in each well, e.g. when aligning the assemblies from column 7 with row E, the assemblies from organism E7 appears in both, allowing identification.

4.2.3.1 Method development: single 8 organism test pool

Before attempting a full 96 organism array, a single test pool of eight unidentified isolates was first processed to optimise the extraction and sequencing and ensure the proposed pipeline was fit for purpose. Unidentified cultures encompassing a range of different morphologies, growth rates, and originally isolated on different media were selected (Figure 4.17). All had been isolated aerobically to make manipulations easier. A pooled sample was created by combining the same amount of liquid culture of all eight organisms. Three different DNA extraction kits along with two different total extraction culture volumes (1.8 and 3.6 mL) were tested to determine which provided the best quality DNA/sequencing data. DNA was also extracted separately from liquid culture derived from each stock to provide a point of reference for the pooled samples. DNA quality control indicated that the different extraction kits and volume of culture extracted had performed similarly.

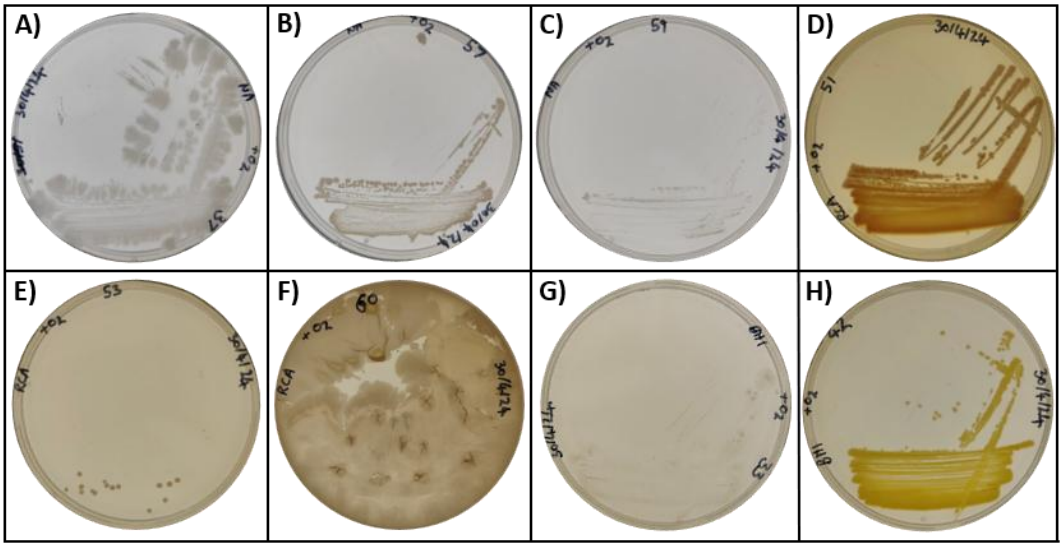


Figure 4.17: Morphologies of organisms A - H.
Images of organisms A – H following streak plating on their respective media types. Organisms A, B, C: nutrient agar; D, E, F: reinforced clostridial agar; G, H: brain heart infusion agar.

16S sequencing was performed on all the pooled and individual DNA samples. Classification at the species level was not achieved with this method but assignment at the genus level was informative (Figure 4.18). Sequencing of samples A – H separately identified A, B, C, E and H as being composed of a single genus. The data suggested that D, F and G contained mixes of up to four different genera, which was surprising given that colonies derived from these stocks had appeared morphologically uniform on plates (Figure 4.17). For the pooled samples, all DNA extraction kits and volumes gave similar results (Figure 4.18) indicating there were no major differences in performance between them. Accounting for duplicates, 10 different genera were identified across the separate samples, all of which were distinguishable in the pooled sample, indicating the pooling had captured all components identified at this level.

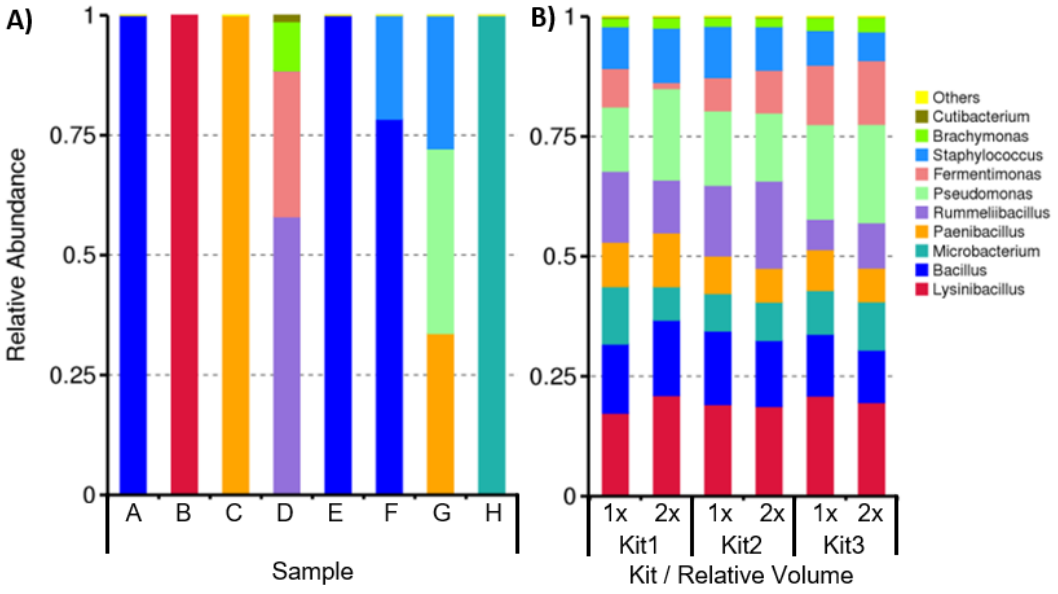


Figure 4.18: Organism A-H identification based on 16S sequencing.
Top 10 most abundant groups at the genus level for organisms A-H assigned based on 16S sequencing. Predicted identity of each group is indicated to the right of the figures. Bar height is proportional to the relative abundance of reads derived from each group from the sample. **A)** Genus level classification of organisms A – H extracted separately using Kit 1. **B**) Genus level classification of the pooled samples containing a mix of organisms A-H extracted using three different kits, each with two different relative extraction volumes. Figures produced by Novogene.

It was not possible to identify organisms to species level from the 16s data, making it challenging to know whether apparently pure samples were a single species belonging to the genus identified or a mixture of species from that genus. These distinctions were

important for the identification effort, so the same pooled and separate DNA samples were submitted for whole genome Illumina short read sequencing. Kraken2 was used to assign identities, initially using unassembled reads. The same five separate samples were identified as likely containing isolates (Figure 4.19, Panel A) in agreement with the 16S data. Classification at the genus level was also in agreement between the two methods, although in most cases the whole genome sequencing provided species level identification (Table 4.3). The same three samples were also identified to be mixes of different organisms (Figure 4.19, Panel B) some of which were classified at the species level (Table 4.3).

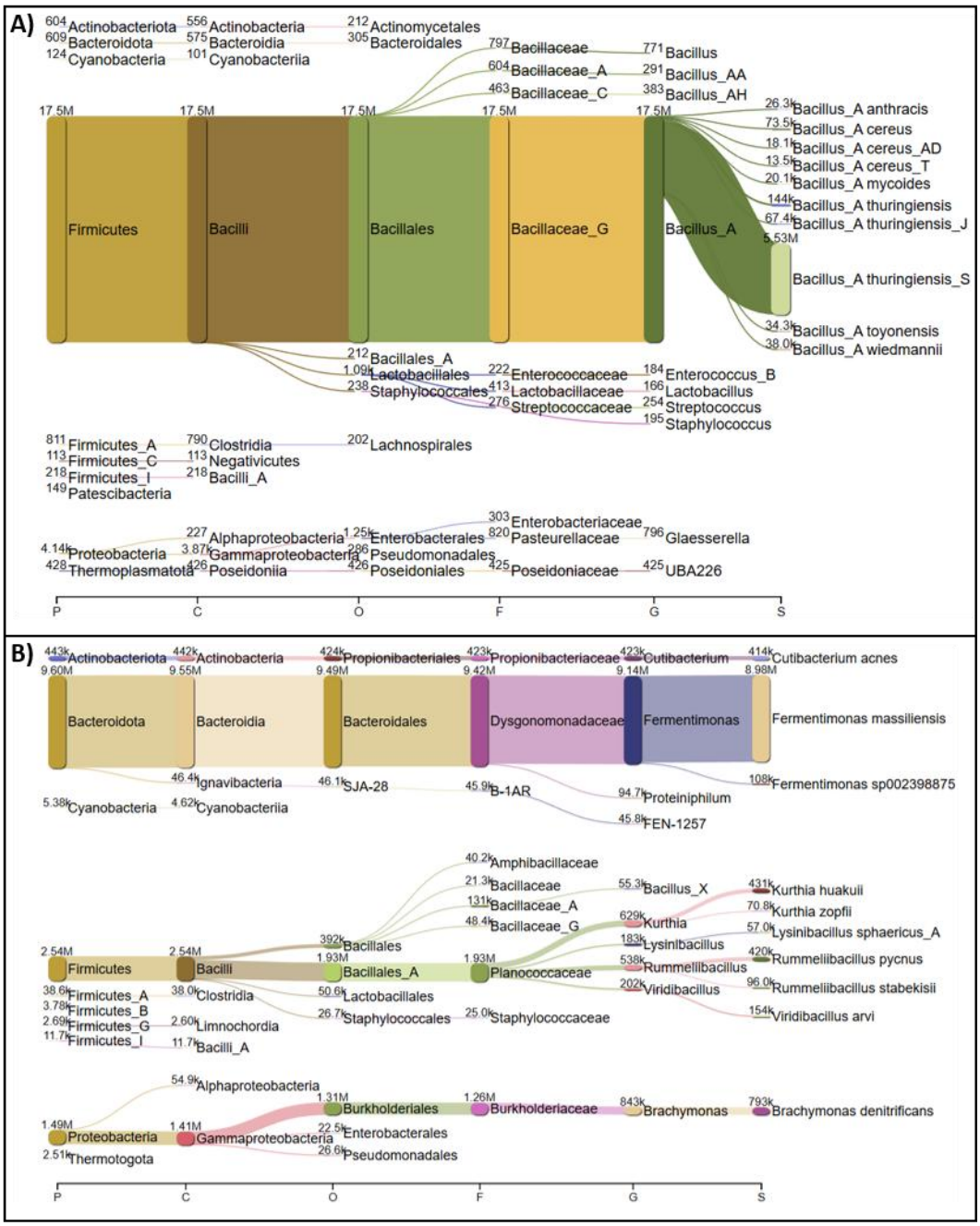


Figure 4.19: Taxonomic classification of unassembled "pure" DNA samples A and D. Examples of Kraken2 taxonomic assignment of unassembled whole genome sequencing reads from DNA derived from **A)** "organism A" (likely an isolate) and **B)** "organism D" (likely a mix). Bars indicate taxonomic assignment of reads from the phylum (left hand side) to the species (right hand side) level (see scale bar at bottom). The number of reads assigned to each classification is indicated above the respective node, with the thickness of the bars giving an indication of the number of reads. Scale bar: P = phylum, C = class, O = order, F = family, G = genus, S = species. Diagrams produced in Pavian.

The pooled samples gave similar outputs to those obtained previously (Figure 4.20) further confirming the apparent lack of impact of DNA extraction protocol on sequencing data.

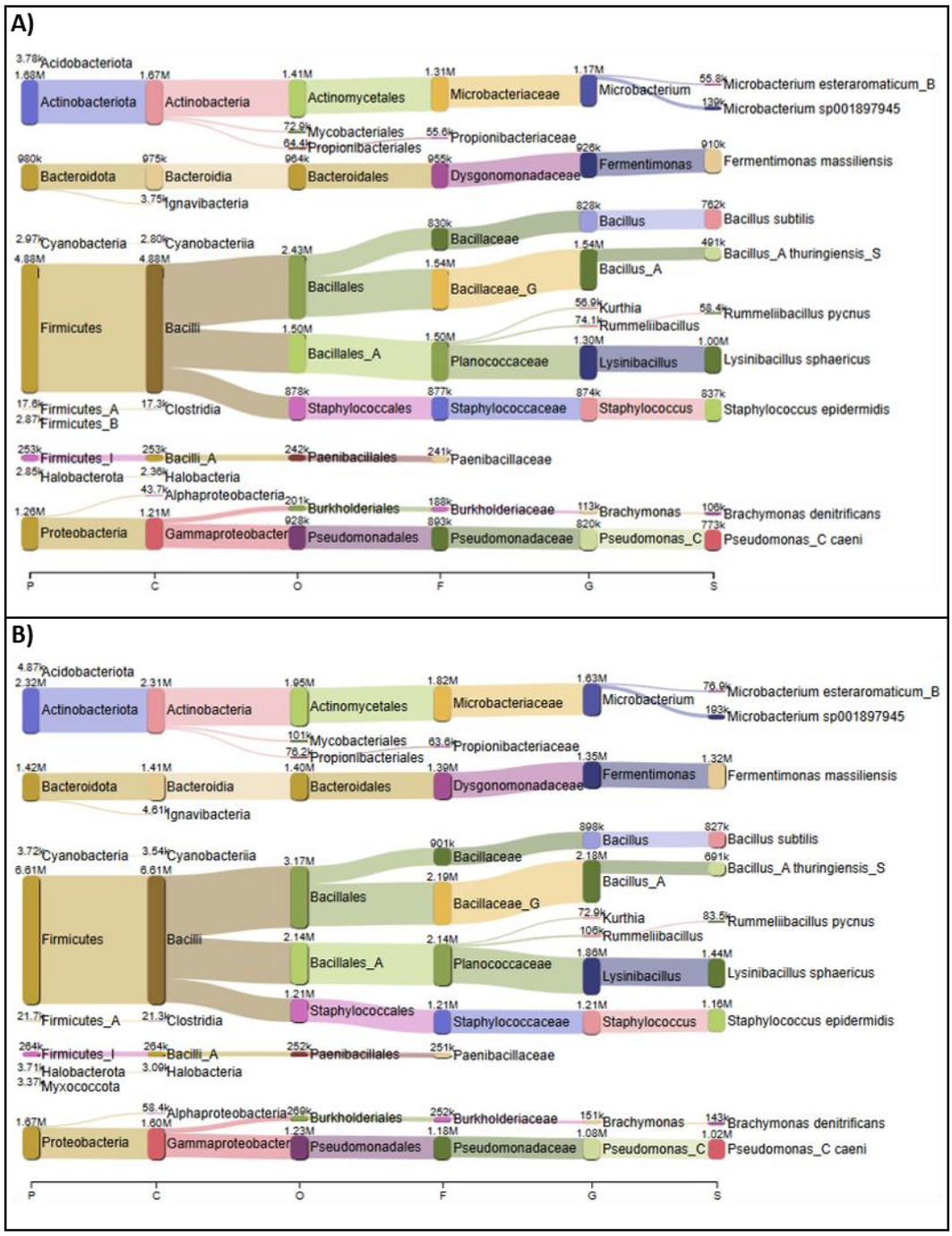


Figure 4.20: Taxonomic classification of unassembled pooled DNA samples.
Kraken2 taxonomic assignment of unassembled whole genome sequencing reads from pooled DNA of "organisms" A – H extracted with different kits using different culture volumes. **A)** kit1 1.8 mL of culture. **B)** kit2 3.6 mL of culture. Bars indicate taxonomic assignment of reads from the phylum (left hand side) to the species (right hand side) level (see scale bar at bottom). The number of reads assigned to each classification is indicated above the respective node, with the thickness of the bars giving an indication of the number of reads. Scale bar: P = phylum, C = class, O = order, F = family, G = genus, S = species. Diagrams produced in Pavian.

As was the case with the 16S sequencing, all organisms identified in the separate samples were also detected in pooled samples (Table 4.3). A crucial distinction between the two methods was that for whole genome sequencing 11 different organisms in total were identified, compared with 10 from the 16S data. This discrepancy was caused by the fact that *Bacillus subtilis* and *Bacillus thuringiensis* which are distinct species belonging to the same genus could not be resolved in the 16S data, highlighting the value of whole genome sequencing for this purpose.

Table 4.3: Summary of contents of unassembled individual DNA samples A-H.
Inferred number/identity of distinct organisms in individual samples A – H. "Number of Organisms" is the number of genuinely different organisms judged to be present in each sample based on analysis of Kraken2 taxonomic assignments of unassembled whole genome sequencing reads. "Distinct organisms" is the cumulative number of distinct organisms present in samples in alphabetical order (i.e. the count only increases if the organism(s) contained in the sample have not been identified previously in others). Brackets after predicted identities indicates taxonomic rank: S = species, G = genus, F = family.

Sample	Number of Organisms	Distinct Organisms (Cumulative)	Identities
A	1	1	*Bacillus thuringiensis* (S)
B	1	2	*Lysinibacillus sphaericus* (S)
C	1	3	*Paenibacillaceae* (F)
D	4	7	*Cutibacterium acnes* (S) *Fermentimonas massiliensis* (S) *Brachymonas denitrificans* (S) *Planococcaceae* (F)
E	1	7	*Bacillus thuringiensis* (S)
F	2	9	*Bacillus subtilis* (S) *Staphylococcus epidermidis* (S)
G	3	10	*Staphylococcus epidermidis* (S) *Pseudomonas caeni* (S) *Paenibacillaceae* (F)
H	1	11	*Microbacterium* (G)

To investigate the whole genome sequencing data further, reads were first assembled and binned, then annotated using the same method as had been used for the unassembled reads. The results for the separate samples were in agreement with those obtained for unassembled reads (Figure 4.21).

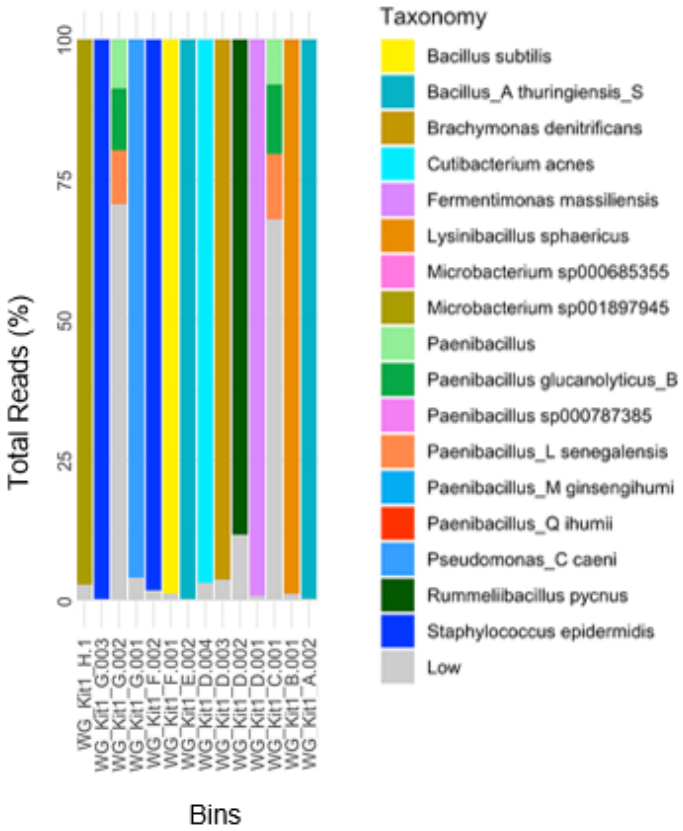


Figure 4.21: Classification of binned individual samples. Kraken2 taxonomic assignment of assembled and binned whole genome sequencing reads from individual DNA of samples A – H at the species level. The sample origin of each bin is provided below the bars. Predicted identity of bin contents is indicated by colour. Bar height is proportional to the amount of assembled sequence in each bin which aligns to the indicated taxonomic assignments.

Similarly, assembled reads indicated all extraction methods tested were comparable, with all component organisms identified in the separate samples also identified in pools (Figure 4.22).

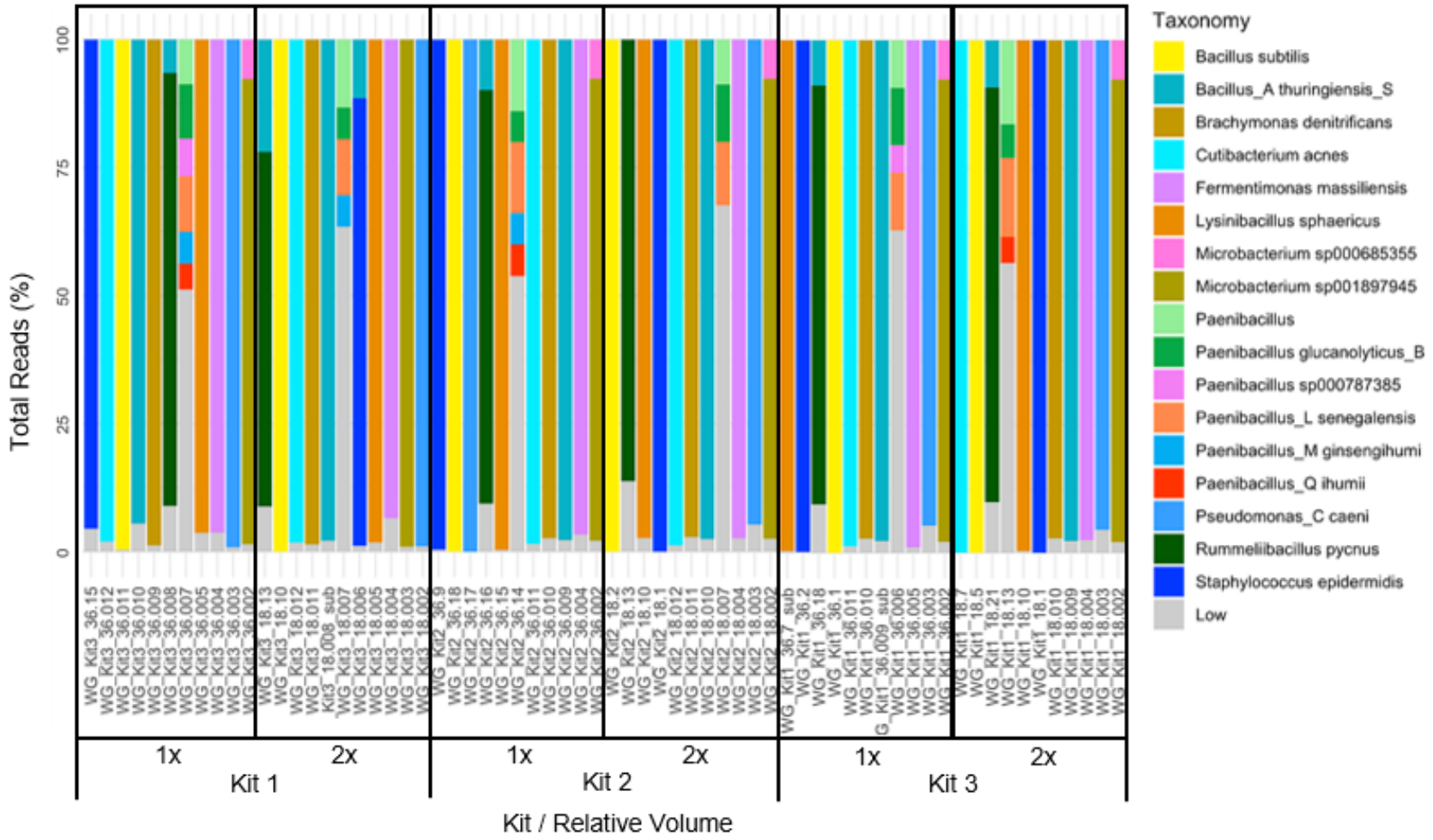


Figure 4.22: Classification of binned pooled samples.

Kraken2 taxonomic assignment of assembled and binned whole genome sequencing reads from pooled DNA of samples A – H extracted with different kits using different culture volumes at the species level. The sample origin of each bin is provided directly below the bars. Predicted identity of bin contents is indicated by colour. Bar height is proportional to the amount of assembled sequence in each bin which aligns to the indicated taxonomic assignments.

Considering the test pool data together this method appeared to have worked well. All component species present in the eight individual DNA samples derived from stocks A – H could be distinguished in the pooled samples, although this had revealed that three of these samples were mixes rather than isolates of single organisms. Whole genome sequencing gave greater taxonomic clarity than 16S, in many cases identifying component organisms to the species level.

4.2.3.2 96 Sample pooling

The organism identification protocol was scaled up to 96 samples (Figure 4.16). No difference in performance between DNA extraction kits or volumes of liquid culture on the quality of sequencing data had been detected in the 8 organism test pool experiment, so the simplest extraction protocol (Kit 1: DNeasy UltraClean Microbal Kit (Qiagen)) was used with the lower pooled culture volume (1.8 mL).

Inoculated cultures were laid out in a 8 x 12 format. This was done as separate 5 mL cultures rather than in a 96 well plate to ensure sufficient material at this method development stage (Figure 4.23). This “mock plate” contained 83 of the colonies previously generated from sludge plating, 34 of which had been isolated and stored anaerobically (Figure 4.23, red boxes) and 49 aerobically (Figure 4.23, green and blue boxes). Among the aerobic samples were the eight organisms (A – H) already identified in the test pool experiment (Figure 4.23, blue boxes).

Six mock inoculations into the three different media types had been prepared alongside the glycerol stock inoculations. All were visibly clear of contamination and were spread throughout the array (Figure 4.23, white boxes). Seven control glycerol stocks which had also been included as part of the reinoculations but had visible signs of growth were also included to determine the identity of this contamination (Figure 4.23, grey boxes). The fact that all mock inoculations included were visibly clear from contamination indicated that setup of this experiment was clean, which suggested that the contamination present in the glycerol stock controls originated in these stocks themselves,

Liquid cultures were pooled in equal volumes vertically and horizontally according to the layout (Figure 4.23) generating 12 pooled columns each composed of 8 distinct samples

and 8 pooled rows each composed of 12. Following DNA extraction, the 20 pooled samples were submitted for whole genome Nanopore long read sequencing to enable subsequent identification.

	1	2	3	4	5	6	7	8	9	10	11	12	
A	Control +	? -	? -	? -	? -	? -	? -	? -	? -	? -	? -	? -	**Row 1 Pool**
B	? -	Control +	? -	? -	? -	? -	? -	? -	? -	? -	? -	? -	**Row 2 Pool**
C	? -	? -	Control +	? -	? -	? -	? -	? -	? -	? -	? -	? -	**Row 3 Pool**
D	? -	? +	? +	Control -	? +	? +	? +	? +	? +	? +	? +	? +	**Row 4 Pool**
E	? +	? +	? +	? +	Control -	? +	? +	? +	? +	? +	? +	G +	**Row 5 Pool**
F	? +	? +	? +	A +	? +	Control -	? +	? +	? +	H +	? +	? +	**Row 6 Pool**
G	? +	? +	? +	? +	? +	? +	D +	? +	E +	? +	? +	? +	**Row 7 Pool**
H	B +	? +	C +	F +	? +	Control +	Control +	Control +	Control +	Control -	Control -	Control -	**Row 8 Pool**
	Column 1 Pool	**Column 2 Pool**	**Column 3 Pool**	**Column 4 Pool**	**Column 5 Pool**	**Column 6 Pool**	**Column 7 Pool**	**Column 8 Pool**	**Column 9 Pool**	**Column 10 Pool**	**Column 11 Pool**	**Column 12 Pool**	

Figure 4.23: Array layout for 96 sample pooling experiment.

Layout of the 96 liquid cultures used for pooling and identification. All cultures in each row and column were mixed to create 20 pools (indicated below and to the right). "?" = unidentified samples, red boxes = anaerobic, green boxes = aerobic, blue boxes = organisms previously identified in the 8 organism test pool. White boxes = uncontaminated mock inoculated controls. Grey boxes = contaminated glycerol controls. + = aerobic conditions, - = anaerobic.

A custom bioinformatics protocol was used to identify the organisms at the intersections of each row and column from the sequencing data. This involved assembling sequences from row and column pools then cross comparing those assemblies. For example, assemblies derived from column 1 were queried against assemblies of each row (1-8), and matching contigs were assigned to the respective cross sections.

Identities were assigned to all cross section bins which matched for more than 10% of their length to an entry in the Kraken2 database. This was cross referenced with the NCBI nucleotide BLAST database and if both assignments were found to agree they were deemed to be identified with “high confidence”. Bins identified by Kraken2 but not BLAST were assigned “intermediate confidence”. Intersections containing either no contigs or bins that did not meet these criteria were classed as not identified, since assignments based on less than 10% of sequence length were deemed to be of little descriptive value.

To give an indication of the quality of identification using this method the results for the intersections corresponding to the previously identified samples A – H were compared (Table 4.4). The identities from the 8 organism pooling experiment were considered more reliable as these had been assigned from separately sequenced samples. Five of the samples were found to contain at least one shared identification between methods. Of the three that did not, sample E was a close match, identified as different species of the *Bacillus* genus. The two way pooling method employed here performed less well for samples G and D, presumably because these were mixes of three or more organisms so computationally more challenging. Despite this, a reasonable match in five of eight cases was seen as a considerable achievement, particularly given that the raw sequencing data used for each method was generated using different sequencing technologies (short read Illumina vs long read Nanopore).

Table 4.4: Comparison of organism identification between different methods.
Inferred number/identity of distinct organisms in individual samples A – H judged to be present based on analysis of Kraken2 taxonomic assignments according to identification of separate samples ("Separate Sample ID") or using the custom bioinformatics pipeline based on intersections of the 96 sample pooling experiment ("Two Way Pooling ID"). Instances where a sample contains at least one component identified as the same by both methods are indicated by ticks.

Sample	Separate Sample ID	Two Way Pooling ID	Agrees?
A	*Bacillus thuringiensis*	*Bacillus thuringiensis*	✓
B	*Lysinibacillus sphaericus*	*Lysinibacillus sphaericus*	✓
C	*Paenibacillaceae*	*Paenibacillus barengoltzii* *Pseudomonas sp001513025* *Bacillus halotolerans*	✓
D	*Cutibacterium acnes* *Fermentimonas massiliensis* *Brachymonas denitrificans* *Planococcaceae*		
E	*Bacillus thuringiensis*	*Bacillus clausii*	
F	*Bacillus subtilis* *Staphylococcus epidermidis*	*Bacillus subtilis* *Kocuria rhizophila*	✓
G	*Staphylococcus epidermidis* *Pseudomonas caeni* *Paenibacillaceae*	*Thiopseudomonas sp001513435* *Gulosibacter molinativorax*	
H	*Microbacterium*	*Microbacterium sp001897945*	✓

Of the 90 samples included in this protocol (excluding blank controls) 56 were putatively identified with high and 10 with intermediate confidence, while 24 could not be identified (Figure 4.24). According to these assignments, 30 of those identified were mixes, and 36 isolates.

The quality of the cultures/DNA appears to have impacted identification, most clearly shown by the results for row 1 in which none of the component cultures were identified (Figure 4.24). While most cultures across the array were visibly turbid after liquid culturing, only three of the 11 cultures in row 1 had turbid growth, one was visibly faint and seven had no visible signs of growth. The concentration of DNA extracted from row 1 was low compared with the other pooled rows (18 ng/µl compared with a mean of 79 ng/µl), and fewer reads were obtained (around 600,000 compared with a mean of

2,000,000 across all pools). The assembly for row 1 was comparatively poorer, with 1109 contigs and an N50 value of around 73,000 bp, compared with a mean of 416 contigs and an N50 of approximately 2,500,000 for the pools collectively.

The potential impact of contamination of these identities should also be considered. Three of the samples derived from contaminated glycerol stocks were identified with high confidence as isolates of *Kocuria rhizophila*, *Clostridium sporogenes* and *Tessaracoccus sp000619965*. Collectively, these organisms were detected in seven of the other unidentified samples. In all but one case these presumptive contaminant organisms were identified in addition to at least one other organism, raising the possibility that they had been introduced alongside the contents of each glycerol stock. The identification of these seven affected samples should therefore be viewed with particular caution.

	1	2	3	4	5	6	7	8	9	10	11	12	
A													A
B	Paenibacillus macerans		Paenibacillus sp002573715	Staphylococcus saprophyticus Enterococcus cecorum	Paenibacillus sp002573715	Acetoanaerobium noterae Enterococcus cecorum		Acetoanaerobium noterae Staphylococcus saprophyticus			Clostridium botulinum	Enterococcus cecorum	B
C	Clostridium perfringens	Bacillus licheniformis Clostridium perfringens Streptococcus equinus		Acetoanaerobium noterae Bacillus licheniformis Clostridium perfringens	Acetoanaerobium noterae Paenibacillus ihumii Sporanaerobacter acetigenes	Paenibacillus fonticola Paenibacillus ihumii	Bacillus licheniformis	Bacillus licheniformis Streptococcus equinus	Veillonella magna	Clostridium botulinum	Clostridium sporogenes Enterococcus cecorum Bifidobacterium thermacidophilum	Clostridium sporogenes	C
D	Clostridium sporenges	Acinetobacter towneri	Advenella sp001513225		Glutamicibacter soli	Corynebacterium efficiens Glutamicibacter soli Staphylococcus warneri		Brachymonas denitrificans	Staphylococcus warneri Bacillus clausii	Clostridium sporogenes Glutamicibacter soli	Advenella sp001513225	Acinetobacter towneri Kocuria rhizophila	D
E	Bacillus pumilus	Gulosibacter sp900163695	Chryseobacterium sp001424105	Gulosibacter molinativorax Pseudomonas caeni			Bacillus licheniformis	Pseudomonas caeni Bacillus subtilis Gulosibacter molinativorax	Micrococcus luteus	Sphingomonas sp003075315	Staphylococcus hominis	Thiopseudomonas sp001513435 Gulosibacter molinativorax	E
F	Paenibacillus naphthalenovorans	Bacillus thuringiensis Massilibacterium senegalense	Lysobacter concretionis	Bacillus thuringiensis	Pseudomonas caeni Tessaracoccus sp000619965			Chryseobacterium sp002308695 Comamonas thiooxydans		Microbacterium sp001897945	Comamonas aquatica Pseudomonas caeni	Candidimonas sp000209655 Tessaracoccus sp000619965	F
G	Bacillus altitudinis	Advenella sp001513225 Hydrogenophaga pseudoflava Thauera humireducens			Bacillus sp001439625	Bacillus licheniformis Bacillus badius Fermentimonas massiliensis Pseudomonas caeni		Enterococcus cecorum	Bacillus clausii	Castellaniella defragrans Petrimonas sp002356435 Pseudomonas caeni	Advenella sp001513225 Bacillus licheniformis Bacillus fordii Castellaniella	Bacillus sp001989355 Pseudomonas caeni	G
H	Lysinibacillus sphaericus	Brachymonas denitrificans Bacillus subtilis	Pseudomonas sp001513025 Bacillus halotolerans Paenibacillus barengoltzii	Bacillus subtilis Kocuria rhizophila	Bacillus subtilis	Staphylococcus pasteuri Pseudomonas sp001513025			Kocuria rhizophila		Clostridium sporogenes	Tessaracoccus sp000619965	H
	1	2	3	4	5	6	7	8	9	10	11	12	

Figure 4.24: Identities for 96 sample pooling experiment.

Identification of organisms in each sample according to bioinformatic analysis of sequence assemblies derived from cross sections of pooled row and column DNA samples. Yellow boxes = organisms identified with intermediate confidence (assembly(s) matching for >10% of their length to an entry in the Kraken2 database). Green boxes = organisms identified with high confidence (criteria as for yellow boxes but identification also corroborated by BLAST). Red boxes = samples with no assemblies or assemblies with poor matches to Kraken2/BLAST. White boxes = intersections derived from controls with contained no assemblies. Thick bordered boxes = previously identified organisms A-H. Identity according to the Kraken2 database is given. Boxes containing multiple identities are mixes.

4.3 Discussion

4.3.1 Aims recap

The objective of this chapter was to investigate methods for obtaining potential AD chassis organisms, either by evolving a lab strain to be fitter in an AD environment or by isolating wild (i.e. AD adapted) versions of laboratory organisms from sludge.

4.3.2 Chapter achievements

The directed evolution experiments were a success in that it was found to be possible to improve the fitness of a laboratory strain of *E. coli* in anaerobic conditions. The isolation of wild organisms with potentially engineerable relatives yielded one target organism (*B. subtilis*). Encouraging progress was also made in devising a cheaper and less labour-intensive microorganism identification protocol based on whole genome sequencing.

4.3.3 Directed evolution

To explore directed evolution as a tool to improve the fitness of a laboratory microorganism for AD, *E. coli* was adapted to be fitter in the absence of oxygen, a condition by definition found in all anaerobic digesters (Harirchi *et al.* 2022). The data suggest that this was successful, as the two best performing *E. coli* populations in the comparative growth experiment had been anaerobically evolved (Figure 4.7). Anaerobic adaptation of the REL4536 strain of *E. coli* has been attributed to mutations in a number of genes including those involved in fermentation (Finn *et al.* 2017). Interestingly, the authors of this study noted that their aerobically evolved populations did not increase in fitness when tested in anaerobic conditions (Finn *et al.* 2017), which while observed in the majority of cases here was not true of population 7 (Figure 4.7).

Whole genome sequencing of DNA derived from the populations evolved in this work revealed possible explanations for some of the observed differences in growth performance. The most interesting variant was a missense mutation in the *frdA* gene (Table 4.1) as it was found only in the two best performing anaerobically evolved populations (Figure 4.7). *frdA* is part of the *frdABCD* operon and codes for the catalytic flavoprotein subunit of the fumarate reductase gene, which is required in anaerobic respiration using fumarate as a terminal electron acceptor in *E. coli* (Maklashina *et al.*

1998; Cecchini *et al.* 2002). One possibility is that this variant improved the enzyme somehow, making cells carrying it in populations 1 and 2 more efficient at anaerobic respiration using fumarate, and thus able to grow faster and to higher final densities.

Interestingly, the status of the allele in populations 1 and 2 is the same as that of WT *E. coli* MG1655 (Table 4.1), suggesting that this is an example of a reversion mutation, whereby a genetic change obtained by bacteria in one environment is costly when they return to their original environment, causing the original version of the gene to be selected for (Avrani *et al.* 2020; Pennings *et al.* 2022). A mutation was presumably introduced at some point during the generation of the ancestral *E. coli* strain MG1655_rLP5(nth-ydgR) used as the basis for the evolution experiments here, which was itself derived from *E. coli* MG1655 (Urtecho *et al.* 2019). This mutation may have been tolerated when grown aerobically where fumarate is not used as an electron acceptor for respiration (Unden *et al.* 2016), but when passaged in anaerobic conditions here cells which underwent a reversion to the original status of the gene may have been selected for. If this were the case, it would be tempting to speculate that the increase in final OD_{600} seen from day three for population 2 during the passage itself could have coincided with the gain of this mutation (Figure 4.6). The corresponding increase only occurred on the final day of passage for population 1, and the fact that the frequency of this variant was 53% in this population (Table 4.1) but near fixation in population 2 (which performed best in the fitness assay) would support this explanation.

The same missense mutation in the *treB* gene, which is involved in trehalose transport (Klein *et al.* 1995) was found in the three best performing anaerobic populations in the fitness assay. The fact that this was the only variant detected in population 3, and that it was close to fixation makes this the most likely candidate for the increased growth phenotype in this population according to the data available. Interestingly, parallel nonsynonymous mutations in this gene were also detected in another long-term evolution study with *E. coli* MG1655 (Behringer *et al.* 2018). How an alternative form of this gene might have provided a growth advantage is not clear, as trehalose would not be expected to be present at appreciable levels in LB medium, where the primary carbon source is amino acids (Sezonov *et al.* 2007).

Population 4 was the worst performing anaerobically evolved population (Figure 4.7) which means genetic differences between it and the other anaerobic populations could be informative. Two missense mutations at relatively low frequencies were present in the *crp* gene in population 4 but absent from anaerobic populations 1-3. *crp* codes for a transcription factor responsible for regulating hundreds of genes in *E. coli* as part of the CRP/cAMP system (Green *et al.* 2014), including many important for carbon metabolism (Shimada *et al.* 2011). In carbon limiting conditions *crp* deletion has been shown to lower growth rate, and loss of a functional CRP/cAMP system renders cells less able to utilise other carbon sources (Franchini *et al.* 2015). Given the central role of this gene in energy metabolism, it seems unlikely the mutations detected in population 4 caused a loss of function as no decline in growth rate relative to the ancestral strain was detected (Figure 4.7). A more likely explanation is that these *crp* mutations are selectively neutral, and it is the absence of variants in the *frdA* and *treB* genes and the possible growth advantages they afford to anaerobic populations 1-3 that explains the comparatively worse performance of population 4.

Population 7 performed the third best well in the fitness assay (Figure 4.7) which was unexpected given that this was conducted in anaerobic conditions, yet this population had been evolved in the presence of oxygen. Three genetic variants were detected, but two were shared with other aerobically evolved populations which did not perform well in the fitness assay, suggesting they were not responsible for the improved performance. One mutation was detected only in population 7 and therefore represents the most likely explanation from the data available. This mutation results in the insertion of a single nucleotide within the annotated regulatory region of the *flhDC* genes (Soutourina *et al.* 1999; Barker Clive *et al.* 2004). These genes act as master regulators of flagellar synthesis: their expression generates a hexameric transcriptional activator (composed of two FlhC and four FlhD subunits) which is at the top of a cascade that positively regulates around 50 flagellar genes involved in *E. coli* motility (Soutourina *et al.* 1999; Wang *et al.* 2006). One possible explanation as to why cells with this variant grew more rapidly and to higher densities is that this mutation in some way disrupts the expression of *flhDC*, leading to reduced or total loss of flagellar expression, which would be expected to save energy as this machinery is known to be costly (Schavemaker and Lynch 2022).

Downregulation of flagellar genes has often been observed in bacterial evolution experiments (Favate *et al.* 2022), and loss of function in motility genes might be expected to be well tolerated in environments where cells are mechanically shaken such as those carried out here. In agreement with this, motile mutants were found to be selected for when motility deficient *E. coli* were grown in non-shaken liquid medium, but this did not occur when the medium was shaken (Parker *et al.* 2019). Additionally, *E. coli* strains rendered non-motile due to deletions in the *flhDC* regulatory region exhibited superior colonisation of mouse intestine, and grew around 20% faster than the motile parent strains on mouse caecal mucus (Leatham *et al.* 2005; Gauger *et al.* 2007). In further experiments, around 90% of *E. coli* MG1655 introduced into mice with a functional *flhDC* regulatory region were found to have become non-motile 15 days after feeding, with all genetically tested colonies having disruptions in the *flhDC* regulatory region (Gauger *et al.* 2007). This indicates a strong selective pressure to lose expression of this operon in an environment where motility does not provide an advantage, which the authors attributed in part to the energy saved being diverted to other functions including growth (Gauger *et al.* 2007). One obvious avenue for further work would therefore be to conduct a motility assay on mutant cells from in population 7, which would be expected to be non-motile.

An explanation based on loss of motility raises further questions. Firstly, if this variant provides a selective advantage in the conditions used here why was the mutation found in only about one third of this population (Table 4.1). One possible explanation is that the mutation had arisen relatively recently so had not yet had sufficient time to reach fixation. Another potential difficulty with this energy saving explanation is that a frameshift in the *flhC* gene was detected in population 8, yet this population did not exhibit the increased growth phenotype seen in population 7. The frameshift impacts over 50% of the gene, including its zinc ion binding domain and regions known to interact with FlhD (Wang *et al.* 2006), so it seems highly unlikely the gene would remain functional. FlhD is able to impact expression of some genes alone without involvement of FlhC (Wang *et al.* 2006), so one possibility is that it was these genes primarily responsible for the phenotype of population 7. Another possibility is the mutation does provide a growth advantage to population 8, but its relatively low frequency of 13.8% meant a high enough proportion of cells did not carry the variant to be clearly detectable in the fitness assay.

Several mutations in other genes were detected in populations which did not perform better in the growth assay. For example, mutations in the *fimA* gene occurred independently at three different locations in three different populations. This is an example of parallel evolution (Bolnick *et al.* 2018) and makes an explanation based on genetic drift (i.e. random chance) seem unlikely. Parallel evolution is observed widely in experimental evolution studies since populations exposed to the same selective pressure often adapt to those conditions in genetically similar ways (Patton *et al.* 2026), so a more likely explanation is that these changes provided a selective advantage. Fimbriae are energetically costly organelles (Müller *et al.* 2009), and *fimA* codes for the major structural subunit of type 1 fimbriae, with around 500 – 3000 copies being required per pilus (Capitani *et al.* 2006). It would be tempting to speculate that the mutations detected here may enable an energy saving, but this does not appear to be the case since two of the populations carrying *fimA* mutations did not exhibit a growth advantage (Figure 4.7). The situation is further complicated by the fact that that type 1 fimbrial genes are phase variable (i.e. exist in a reversible but genetically determined "on" of "off" state) in *E. coli* (Kulasekara and Blomfield 1999), and the proportion of a population in either state can vary considerably based on many different factors (van der Woude and Bäumler 2004). The reasoning behind this parallel evolution remains unclear: perhaps these variants provided some kind of selective advantage not detected in the experiments carried out here.

While some of these associations between genetic variants and growth phenotypes seem compelling, a current limitation is that it is only possible to report correlations between the two with the existing data. It is not possible to determine solely from sequencing data whether a missense mutation leads to a gain or loss of function (Behringer *et al.* 2018), and one possible avenue for further work would be to test for causality by reintroducing the most interesting mutations detected here into an otherwise isogenic strain of *E. coli*, for example using the well established λ Red system (Murphy 2016).

Another limitation to the directed evolution carried out in this chapter is that the results give no indication as to how well the evolved strains would perform in an actual anaerobic digester. In reality there are likely to be many other factors beyond the anaerobic environment itself which exert a selective pressure on organisms in this environment,

such as competition with endogenous organisms or ability to utilise energy sources in the feed. Testing whether the strains generated here were fitter among an AD community and investigating persistence and stability in this environment would be excellent candidates for further work.

Importantly the RFP expressing strain used here is clearly distinguishable even from WT *E. coli* on solid media plates (Figure 4.3) so a test involving inoculation of this strain into vessels containing anaerobic sludge and estimation of survival rates over time by plating would be straightforward. Significantly higher colony counts of the best performing populations compared with the ancestral control would suggest that adaptation to the anaerobic environment alone was sufficient to increase fitness within a real AD community. Even if this were found to not be the case, further rounds of experimental evolution to other AD relevant parameters could be iteratively carried out on the best performers from previous directed evolution efforts in an attempt to generate an even more competitive organism.

4.3.4 Selective plating of sludge

The second approach in the chassis organism portion of this chapter was to attempt to isolate close relatives of laboratory organisms from AD, based on the assumption that these would be fitter in the environment from which they were isolated than their laboratory adapted relatives. While approaches used to identify candidate organisms to search for from sequencing data had limited success, three organisms whose laboratory counterparts have well developed genetic tools (*C beijerinckii* (Patinios *et al.* 2023), *E. coli* (Pontrelli *et al.* 2018) and *B subtilis* (Errington and Aart 2020)) were selected as targets.

Theoretically, if an organism is sufficiently well characterised it should be possible to isolate it out from a mixed culture using highly selective plating approaches tailored specifically to that organism. Identifying validated isolation protocols with this level of specificity was found to be unexpectedly challenging, possibly because existing well established approaches seemed to primarily be restricted to microorganisms of clinical relevance (Gasanov *et al.* 2005; Park *et al.* 2012; Law *et al.* 2015). A compromise involved adapting a series of previously validated methods which were not specific to the target species but were selective for a group to which each belonged (coliforms for *E. coli* (The

Standing Committee of Analysts 2016), sulphite-reducing clostridia and *Clostridium perfringens* for *C. beijerinckii* (The Standing Committee of Analysts 2021), and aerobic spore-forming bacteria for *B. subtilis* (Environment Agency 2007)). 17 organisms isolated from this selective plating effort were sequenced, 10 of which were identified down to the species level through whole genome sequencing.

The majority of species recovered on MLGA plates were identified as *Pseudomonas guangdongensis*, which was previously isolated from a microbial fuel cell (Yang *et al.* 2013). One sample contained *Alcaligenes faecalis*, which has been investigated as a source of potentially medically useful metabolites (El-Sayed *et al.* 2024). Both samples identified to species level on TSL plates were *Paraclostridium masiliensis*, which was described recently following isolation from human stool (Bellali *et al.* 2019). One of the organisms isolated from a NAB pate was identified as *B. subtilis*, i.e. the organism this plate type had been selected to attempt to obtain on account of its well developed genetic tools (Errington and Aart 2020). *Bacillus lichenformis,* for which genetic tools are also available (Yang *et al.* 2025) was also identified from this plate type.

There was insufficient time to take this further, but continued work in this area would involve attempting to genetically manipulate promising successfully isolated organisms, since the experimental approach has been designed such that it should be possible to retrofit existing tools from laboratory equivalents into the target species. An obvious starting point would be to attempt existing transformation methods to introduce DNA, for example in the case of *B. subtilis* established protoplast based (Chang and Cohen 1979) or electrocompetent (Xue *et al.* 1999) protocols could be tested to transfer in a plasmid carrying an easily screenable marker like GFP.

It should be acknowledged that the tool transfer process from laboratory strains to wild relatives may not be trivial. For example, *Clostridium* are known to generally be challenging to transform on account of their restriction modification systems (Dong *et al.* 2010). It is anticipated that the existing literature should in most cases provide potential solutions. For example, transformation issues related to RM systems can be overcome by deleting said systems (Foulquier *et al.* 2019), pre-methylating DNA in an intermediate host

(Liu and Deutschbauer 2018), or transmitting DNA into the host through conjugation (Woods *et al.* 2019).

An alternative option would be methods such as *in situ* transfer in which an engineered donor bacterium is used to transfer a genetic cargo to endogenous community members in their native environment through horizontal gene transfer (Ronda *et al.* 2019). This overcomes some of the potential pitfalls with the existing approach since selectable markers can be transferred into existing community members (which could subsequently be isolated out) without the need for prior knowledge of the organisms within the community (Wang *et al.* 2019; Rubin *et al.* 2022; Kalvapalle *et al.* 2025).

4.3.5 Non-selective plating of sludge

Untargeted plating and subsequent storage were used to isolate 95 organisms across three different media types, 34 anaerobically and 61 aerobically. Of these the majority were found to be recoverable from glycerol stocks (Figure 4.15), indicating high stability in cryostorage.

Results indicated a possibly higher proportional rate of recovery for aerobic isolates, but this should be viewed with caution as a higher proportion of the anaerobically isolated colonies did not show visible growth when their respective glycerol stocks were being prepared. Preliminary plating work indicated that this does not necessarily mean organisms were not present and recoverable, but it is possible some proportion of these anaerobic isolates were not viable at the point of storage. It can be concluded at this point that the outgrowth, storage and revival process collectively had a lower success rate for anaerobic colonies, but precisely what part of this process was responsible cannot be determined and would be a possible avenue for further work.

Length of time in cryostorage had no impact on recovery success rates for the subset of aerobic organisms tested. Even though recovery was tested up to a period of 13 months, possible loss of organisms in storage over longer durations cannot be ruled out. How long a bacterial culture can remain viable in cryostorage remains an open question with some studies reporting successful revival after more than 10 (Kaijalainen and Palmu 2015) or even 20 years (Hare *et al.* 2015). Stability in storage is likely to depend on a variety of

factors such as the organism itself and method of storage (Berner and Viernstein 2006; Senaratne and Jayaweera 2024) so may not be easily predictable within a given system without actual testing.

These plating experiments were found to be extremely labour intensive. For example, the experiment in which three different media types were compared anaerobically and aerobically required the preparation of over 200 solid medium plates followed by manual picking of colonies and was likely at the limit of what a single person can process at one time. This is problematic because is constrains the number of conditions that can be tested fairly, as a second experiment conducted later would have to be conducted on a different sludge sample, either isolated later or stored for some period of time, making comparisons between the two unreliable. Further conditions that might have been interesting to test but were not possible due to this included additional media types or plating in aerobic conditions followed by incubation in the anaerobic chamber, which may have provided a solution to some of the contamination issues that were encountered for the anaerobically isolated samples.

4.3.6 Sequence based microorganism identification

A method suitable for identifying organisms isolated in this project using the resources and expertise available was sought. Whole genome sequencing or individual isolates would have been prohibitively expensive, so an approach based on whole genome sequencing of pooled samples was attempted.

Initial work using an 8 organism test pool suggested the pooling approach performed well. Results based on 16S and whole genome sequencing (both with and without assembly) were broadly in agreement and capable of detecting all components identified in separate samples in the pools. Better taxonomic resolution was achieved using whole genome sequencing, as would be expected given that this approach is able to use substantially more genetic information for identification than 16S methods (Ranjan *et al.* 2016).

An unexpected finding was that several of the glycerol stocks originally thought to be pure isolates were in fact mixes of different organisms. Some simple mixes of perhaps two organisms had been anticipated, but the high apparent proportion of mixes among the

samples (3/8 tested using this method) and the complexity of these mixes (in one case up to four separate organisms) was surprising, particularly given that colony morphologies on plates were indistinguishable including for this sample (Figure 4.17). Potentially even more surprising is that two samples (A and E) originally thought to be distinct organisms based on plate morphology (Figure 4.17) were found to both be an isolate of the same species through sequencing. Images of morphologies similar to those obtained here for *Bacillus thuringiensis* in sample A (Park *et al.* 2017) and E (Tao *et al.* 2014) could be found in the literature, suggesting this organism can appear very different on plates. This highlights the limitations of using colony morphology as a guide for microorganism identification.

Some challenges were encountered during the process of scaleup to a 96 sample format. This approach did provide a putative identification for the majority of organisms included, and in many cases agreed with identifications made with higher confidence from separately sequenced samples at the test pool stage. However, the method appeared to struggle with cultures containing mixes, and duplicates (i.e. the same organism(s) appearing multiple times in a row or column) were also challenging to resolve. It is possible to generate pure cultures from mixes on plates by restreaking (Sanders 2012). While this was deliberately avoided in this work to try to preserve possible symbiotic interactions (which are known to be widespread in AD (Saha *et al.* 2020)), this would be a worthwhile amendment to attempt as further work, as it is anticipated based on the data obtained that if all of the cultures in the 96 sample experiment had truly been isolates the method would have performed much better.

4.3.7 Next steps

In this chapter methods for obtaining better chassis for AD were investigated, with the idea that these could serve as a basis for strain engineering protocols to introduce novel functions to AD (as explored in Chapter 3). Preliminary results suggest this has some promise, but more work would need to be done to develop these approaches to a point where chassis which might be useful in AD can actually be obtained. Yet genetic engineering of microbial components of the AD system is just one way through which performance improvements might be biologically achieved. Another possibility would be to control the naturally existing composition of the wild community itself.

Chapter 5: Culturomics

5.1 Introduction

5.1.1 Context of work

The overall aim of this PhD project is to investigate biological approaches for monitoring and manipulating microbial communities in AD. Another possible approach to achieving this is to manipulate the composition of the AD community as a whole. This differs from the approaches previously used in the project because it is not reliant on manipulation of individual organisms on the genetic level. An automated colony picker had become available to the lab by this point which was anticipated to make this approach more feasible.

5.1.2 Possible advantages of a simplified AD community

While AD communities tend to differ substantially in overall composition, functional guilds (i.e. groups of organisms performing the same key functions) remain constant (Vanwonterghem *et al.* 2016; Lim *et al.* 2020; Dueholm *et al.* 2024; Keating *et al.* 2025). This leads to the intriguing possibility that AD could theoretically be performed by a simplified community represented by one or a few members of each of these functional guilds. A simplified AD community might be expected to be less resilient to fluctuations in conditions, as it is functional redundancy in the system that is thought to be responsible for this (De Vrieze *et al.* 2017; Lim *et al.* 2020; Cao *et al.* 2025; Jourdain and Gu 2025). However, if a highly simplified AD community could be produced it may be expected to provide improved predictability compared to more complex "wild" communities, which could translate into increased control over the AD process. Such a community would also be of high scientific value as it would enable exploratory experiments like controlled community member dropouts (Jourdain and Gu 2025).

5.1.3 An AD SynCom as a simplified community

One way in which a simplified AD community could be generated would be to create a synthetic mixed AD community or SynCom (see section 1.23.2, General introduction). Generating such a community using a bottom-up approach (i.e. assembly from individual defined microbial isolates (Lawson *et al.* 2019)) would come with the advantage that any

resulting community would be completely defined and therefore theoretically easier to study and control (Cheng *et al.* 2022), so will be the approach investigated here. An essential requirement for producing such a SynCom is to isolate out and collect strains from the environment of interest (Northen *et al.* 2024).

5.1.4 Culturomics

While isolation of microorganisms from mixed samples has been practiced for well over 100 years (Bonnet *et al.* 2020), "culturomics" which involves high throughput culturing (Howe *et al.* 2024) using techniques like mechanisation and miniaturisation (Huang *et al.* 2023) has emerged as an alternative to traditional techniques (Wang *et al.* 2024). Low throughput isolation methods such as those used in Chapter 4 are labour intensive and slow (Huang *et al.* 2023). With a colony picker it should be possible to increase the throughput by applying culturomics principles to AD communities. An important consideration is which community/communities should be used to source these isolates.

5.1.5 Working from consistent communities

Ideally it would be possible to do all experimental work from a single static microbial community such that all results are truly comparable. As was discovered in the previous chapter there is a limited amount of work that can be carried out from a single sample taken from an AD system. This necessitates further sampling which is undesirable since AD communities vary over time and space (Calusinska *et al.* 2018). The THP pretreated sludge used throughout this project had been selected over non-pretreated material in consideration of this issue. This is based on the principle that since AD is a closed system, the main source of ingress of new microbial community members in a running digester is with the feed (Fernandez-Gonzalez *et al.* 2021). The heat and pressure during THP treatment is thought to sterilise the incoming feed (Barber 2016) so digesters running on this material theoretically lack a major potential source of microbial species influx (a form of dispersal in ecological terms (Nemergut *et al.* 2013)) that would be present in systems in which feed is not heat pretreated. However, since other ecological mechanisms such as selection, drift and speciation also act on digester microbial communities (Vellend 2010; Nemergut *et al.* 2013), composition would still be expected to change over time even in the absence of dispersal. For this reason some means of obtaining a temporal snapshot of the mixed microbial communities through storage could provide a potential

means of retaining a more consistent test system (Coker *et al.* 2022; Jourdain and Gu 2025) that can be used over a longer period of time.

5.1.6 Storing mixed microbial communities

Methods for storing mixed microbial communities are less established than those available for axenic cultures (Kerckhof *et al.* 2014), and research in this area specific to AD is limited (Hagen *et al.* 2015). While some studies have explored different methods such as freezing (Hagen *et al.* 2015) drying (Bhattad *et al.* 2017) and use of protective chemicals (Yarberry *et al.* 2019) for AD community preservation, it is likely that different methods are more suitable for different inocula (Castro *et al.* 2002). For this reason, it would be worthwhile to determine whether a viable community can be recovered from a stored sample of Esholt sludge (i.e. the material used throughout this project). For the purposes of this test, "functional" is defined as a microbial community which is able to carry out AD: i.e. convert complex organic molecules into the terminal AD product methane (Wainaina *et al.* 2019). Freezing (i.e. storage of samples at -20^{o}C) will be used as it is a straightforward method that does not require any specialist equipment or reagents, so is likely more applicable in an industrial setting.

5.1.7 Plating of anaerobes in aerobic conditions

The colony picker available was located in aerobic conditions, which is clearly not ideal given the study system is AD. Methods for creating specialist solid media for propagation of organisms which are difficult to cultivate have been reported (Tanaka *et al.* 2014; Kawasaki and Kamagata 2017). These are based on the principle of autoclaving agar and phosphate media components separately, limiting the production of potentially harmful products such as reactive oxygen species (ROS) (Tanaka *et al.* 2014; Kawasaki and Kamagata 2017), and creating an environment supposedly more permissive for difficult to culture organisms (Tanaka *et al.* 2014). This approach was taken further by adding a mix of antioxidants to a different separately mixed base solution, and was reported to be suitable for isolation of strict anaerobes under aerobic conditions (Dione *et al.* 2016). These methods could potentially also overcome the contamination issues which were observed for some experiments in the anaerobic chamber in Chapter 4, since plate manipulations on these specialist plates could be performed under sterile conditions (i.e. in a laminar flow hood) outside of the chamber.

5.1.8 Aims

The overall aim of this chapter is to carry out proof of principle experiments towards generating an AD SynCom. This will involve:

1. Determine whether a functional AD community from Esholt digestate can be stored by simple freezing and subsequently be revived.
2. Develop a method to isolate, culture and store AD organisms at higher throughput using an automated colony picker.

5.2 Results

5.2.1 Investigating system feed

The Esholt site had been selected to source samples for this project because it was thought it would have a simpler community composition than a conventional AD site due to THP pre-treatment killing microorganisms in the feed, removing this source of potentially new organisms for the community. This reasoning relied heavily on the assumption that THP sterilises sludge, however while this has been reported in the literature experimental evidence to support this based on direct viability testing was found to be unexpectedly limited (see section 5.3.3, Discussion). The decision was therefore taken to investigate the impacts of THP in the system used throughout this project.

5.2.1.1 Plating Esholt pre- and post- THP feed

If THP pre-treatment kills microorganisms present in the feed then there should be no (or fewer) organisms present post-THP compared with pre-THP. To investigate this pre- and post- THP Esholt feed was plated on a simple growth medium and colonies counted (Figure 5.1). While mean colony forming units (CFU) were slightly lower from post-THP samples (5.78×10^7 +/- SE 0.39×10^7) than pre-THP samples (8.0×10^7 +/- SE 1.11×10^7) there was no significant difference between the two feed types (t-test: $t = 1.886$; d.f. = 4; $p = 0.1323$) which was unexpected.

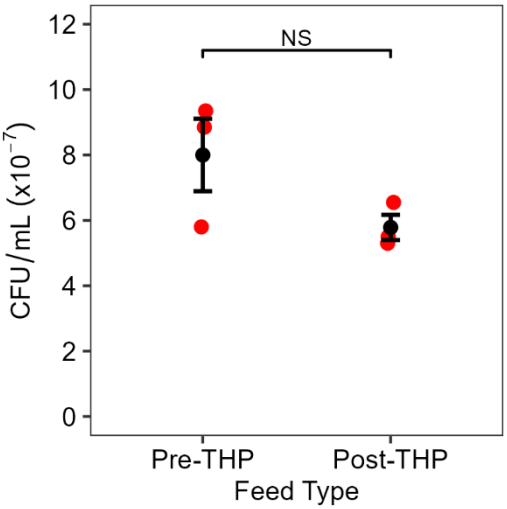


Figure 5.1: **Colony counts of pre- and post-THP plated feed.**
Mean colony forming units per mL of wet feed (black points) calculated from triplicate plate counts (red points) of both feed types. Error bars are +/- one standard error. There was no significant difference between Pre-THP (8.0×10^7 +/- 1.11×10^7) and Post-THP (5.78×10^7 +/- 0.39×10^7) feed CFU; t = 1.886; d.f. = 4; p = 0.1323.

Visual inspection suggested that the diversity of organisms on the post-THP plates might be lower than those on the pre-THP since they were morphologically more similar (Figure 5.2). Post-THP colonies were also smaller. One possible explanation was that the post-THP community was less diverse since only physically robust microbial species able to withstand the aggressive pre-treatment (for example spore formers) were able to survive.

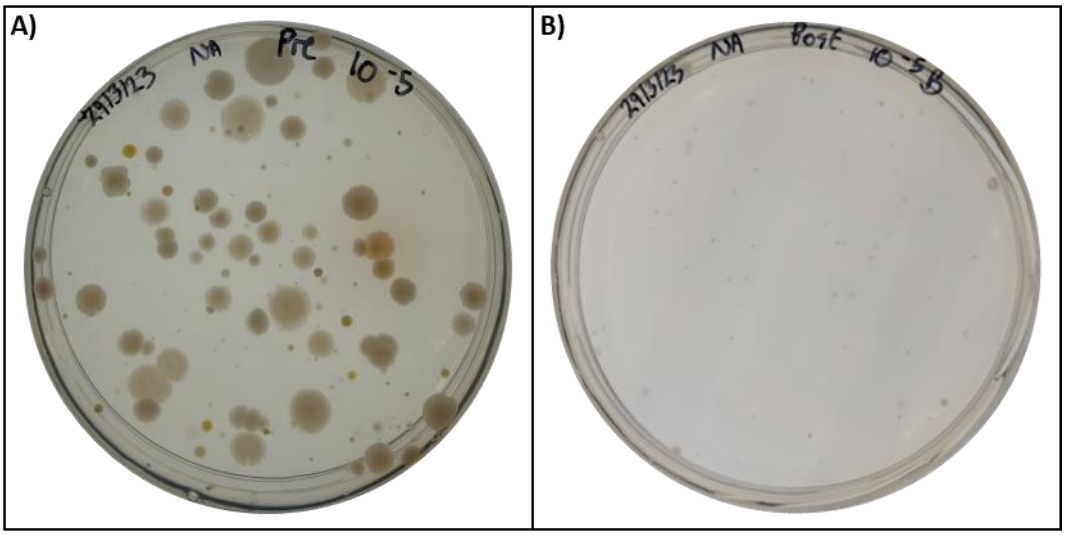


Figure 5.2: Colony morphologies of pre- and post-THP plated feed.
Representative images of both feed types plated on nutrient agar. **A)** Pre-THP feed. **B)** Post-THP feed.

5.2.1.2 16S DNA analysis of pre- and post- THP feed

To investigate possible differences in diversity between pre- and post-THP communities, DNA samples which had been extracted from feed used for the plating experiment (Figure 5.2) were analysed through 16S sequencing. A digester DNA sample obtained and extracted on the same day was also included as a point of comparison. Rarefaction curves indicated that this sequencing captured the majority of diversity in these communities (Figure 5.3, Panel A). Diversity as estimated by number of ASV's (Figure 5.3, Panel B) indicated that the pre-THP feed community had a substantially higher number of unique ASVs (1518) compared with the post-THP feed (78). The number for the digester community was intermediate between the two (647).

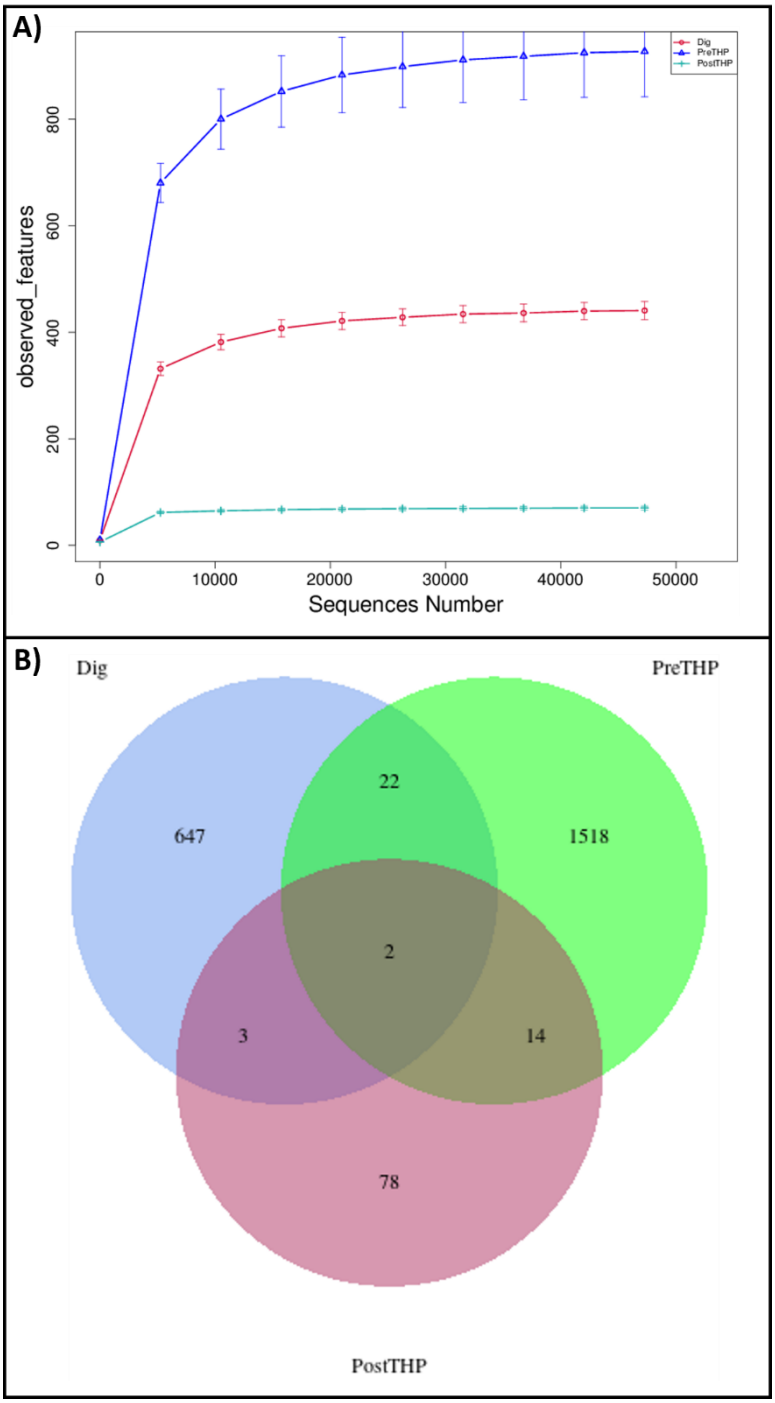


Figure 5.3: Number of ASVs in different conditions.

A) Rarefaction curves for triplicate technical replicates of observed features from different DNA sample sources: blue = pre-THP; red = digestate; green = post-THP. **B)** Sequence features (numbers) shared between (overlapping sectors) and unique to (outside sectors) each treatment. Values are means of three technical replicate 16S sequencing reactions carried out on DNA derived from each condition. Figures produced by Novogene.

Alpha diversity metrics for the three sample types (Table 5.1) further supported these conclusions. The Shannon diversity index, which accounts for both total number and evenness of species and therefore gives a more comprehensive idea of diversity than total number of ASVs indicated the highest diversity in pre-THP samples and lowest in post-THP. Pielou's index, which reports only on species evenness (irrespective of total number of species) was similar for pre-THP feed and digestate but much lower for post-THP feed, indicating that this community was more uneven (i.e. tending more towards dominance of a smaller number of species).

Table 5.1: Alpha diversity metrics for different conditions.
Values are means of three technical replicate 16S sequencing reactions carried out on DNA derived from each condition.

Condition	Shannon Index	Pielou_e
Digestate	6.67	0.76
Pre-THP	7.75	0.79
Post-THP	3.65	0.60

Relative abundance estimates for the 10 most abundant groups found at different taxonomic levels from the 16S data suggested a similar trend. At the class level (Figure 5.4, Panel A) there was more diversity among the 10 most abundant organisms in the pre-THP community. Post-THP communities were primarily composed of the classes *Bacilli*, *Clostridia* and *Negativicutes*, all of which contain representatives capable of forming spores (Yutin and Galperin 2013) which could explain their dominance. The archaeal class *Methanosarcinia* was found only among the top 10 most abundant groups in digester samples which was unsurprising given that like all methanogens these are strict anaerobes (Buan 2018; Blasco *et al.* 2022). Large portions of the community were also composed of *Bacteroidia* and *Clostridia*, both known to be involved in AD (Lee *et al.* 2017). The large relative size of the "other" category, which is composed of all groups less abundant than those shown, indicates that a large proportion of the digester community was composed of rarer groups.

At the genus level (Figure 5.4, Panel B) both the pre-THP and digester communities were dominated by the "other" group which would be expected given that these samples had

appeared to be more diverse according to the alpha diversity metrics (Table 5.1). The most abundant genera in the digester sample included *Methanosaeta*, *Bacteroidetes*, *Sedimentibacter, Rikenellaceae* and *Christensenellaceae* which have all previously been reported in AD (Walter *et al.* 2019; Nakasaki *et al.* 2020; Kieft *et al.* 2023). Two abundant genera found in the pre-THP community (*Peptostreptococcus* and *Lactobacillus*) were also found at increased proportional levels in post-THP feed indicating they may have originated prior to THP and survived THP treatment. Surprisingly, none of the three genera found to dominate the post-THP community are known to be capable of forming spores (Maki and Looft 2018; Rossi *et al.* 2019; Srinivasan *et al.* 2019; Verdecia *et al.* 2019).

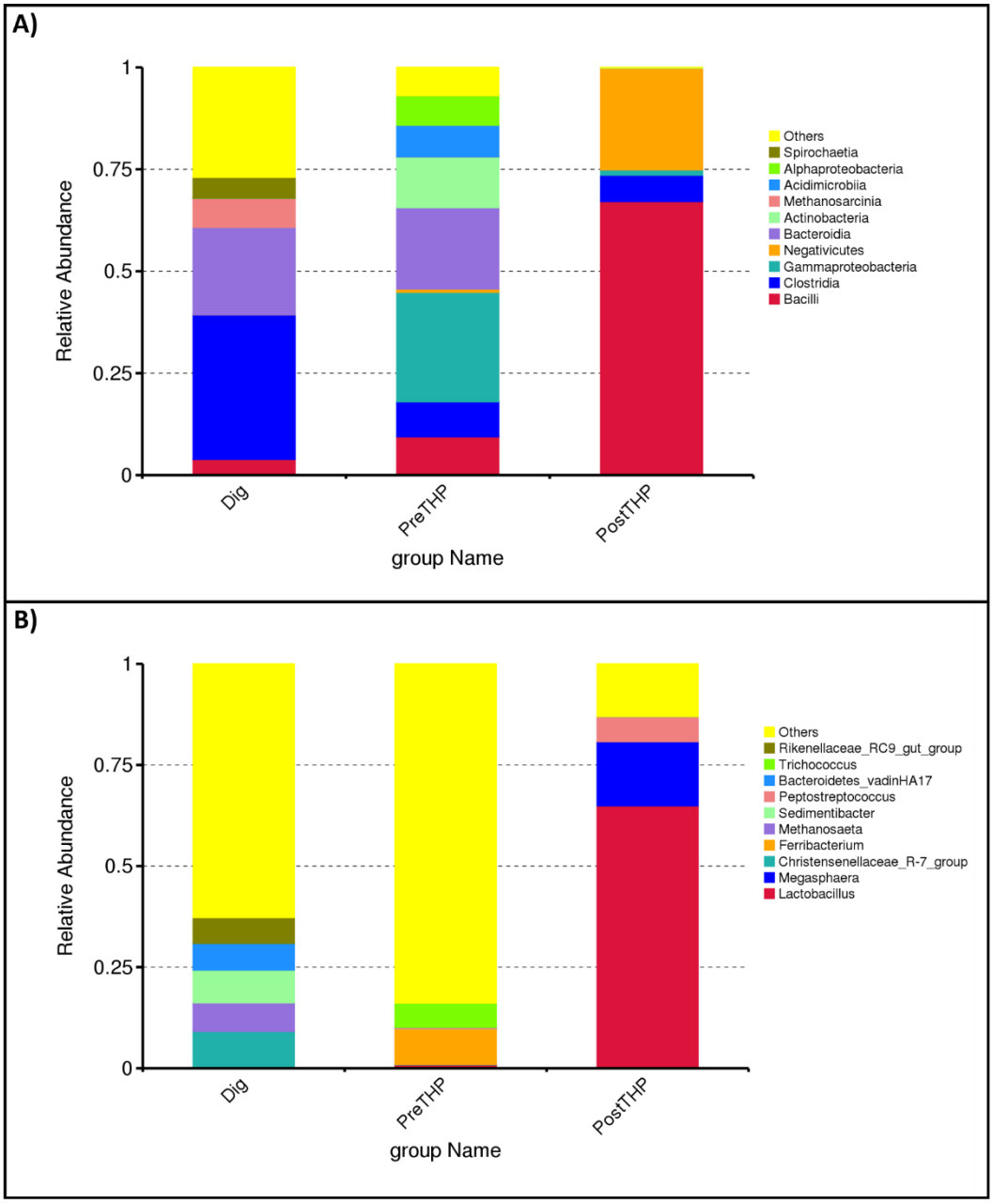


Figure 5.4: Taxonomic classification of abundant groups between conditions.
Top 10 most abundant groups across samples at the **A)** class and **B)** the genus taxonomic levels. Predicted identity of each group is indicated to the right of the figures. Bar height is proportional to the relative abundance of reads derived from each group from the sample. Bars are averages of three technical replicates. Figures produced by Novogene.

5.2.1.3 DNA analysis of pre- and post- THP plate washes

The 16S analysis (section 5.2.1.2) was performed on DNA extracted from the same feed samples used to produce the plates (Figure 5.2) but would have been composed of all DNA present in this community including from organisms that did not grow on these plates. To determine the identity of the organisms specifically present on the plates themselves, DNA was extracted from plate washes. DNA yields from standard extractions were very low, and troubleshooting using an alternative method tailored for isolating DNA from spore forming organisms failed to improve this.

Insufficient DNA was obtained for sequencing with the commercial provider used previously, so the samples were instead 16S sequenced following an additional DNA amplification step. Analysis of the results suggested that this had failed to produce biologically meaningful data since the diversity estimated for a control platewash containing no visible colonies vastly exceeded that of plates which had contained numerous visible colonies. This was likely due to noise/misamplification resulting from the higher number of PCR cycles required for library preparation with the low quantity of input DNA.

Taken together, this investigation had revealed that the THP treated Esholt feed sample tested here was not sterile. While THP treatment did appear to lower the diversity of microorganisms, the fact it contained a similar number of organisms to untreated feed was unexpected and further justified the potential value of investigating approaches to store an AD community to prevent the need for repeated digester sampling.

5.2.2 Can a viable AD community be stored?

A biological methane potential (BMP) test was carried out to determine whether an AD community derived from the material used throughout this project would remain viable (i.e. still produce methane) following storage by freezing at -20 ^{o}C. To provide starting material for the BMP test, fresh Esholt sludge was starved until no further biogas production was detected (Starved, Figure 5.5). A portion was then frozen and thawed (Frozen, Figure 5.5) and a second control sample was not (Not-Frozen, Figure 5.5).

For the test, starved sludge which had been frozen and thawed was split into two treatment groups: Frozen Fed (Figure 5.5) was provided with biologically degradable organic material (cellulose) and Frozen Unfed (Figure 5.5) continued to receive no feed. A positive control consisting of starved sludge which had not been frozen, either fed (Not-Frozen Fed, Figure 5.5) or not fed (Not-Frozen Unfed, Figure 5.5) was also included.

Since a complete AD community is required to produce methane, production of methane in the Frozen Fed treatment would demonstrate that a viable AD community had survived freezing, whereas no methane production would suggest a functional community was no longer present. DNA samples were taken at different stages of the experiment to detect changes in community composition as a result of different treatments.

Alongside this experiment, 320 two mL aliquots of the remaining starved sludge (Starved, Figure 5.5) were stored at -20 ^{o}C. Since these were subjected to the same treatment as the frozen samples in the BMP test (Frozen, Figure 5.5), if those samples could be demonstrated to be viable following thawing (Frozen Fed, Figure 5.5) then these 320 aliquots could serve as common stored source material for experiments going forward.

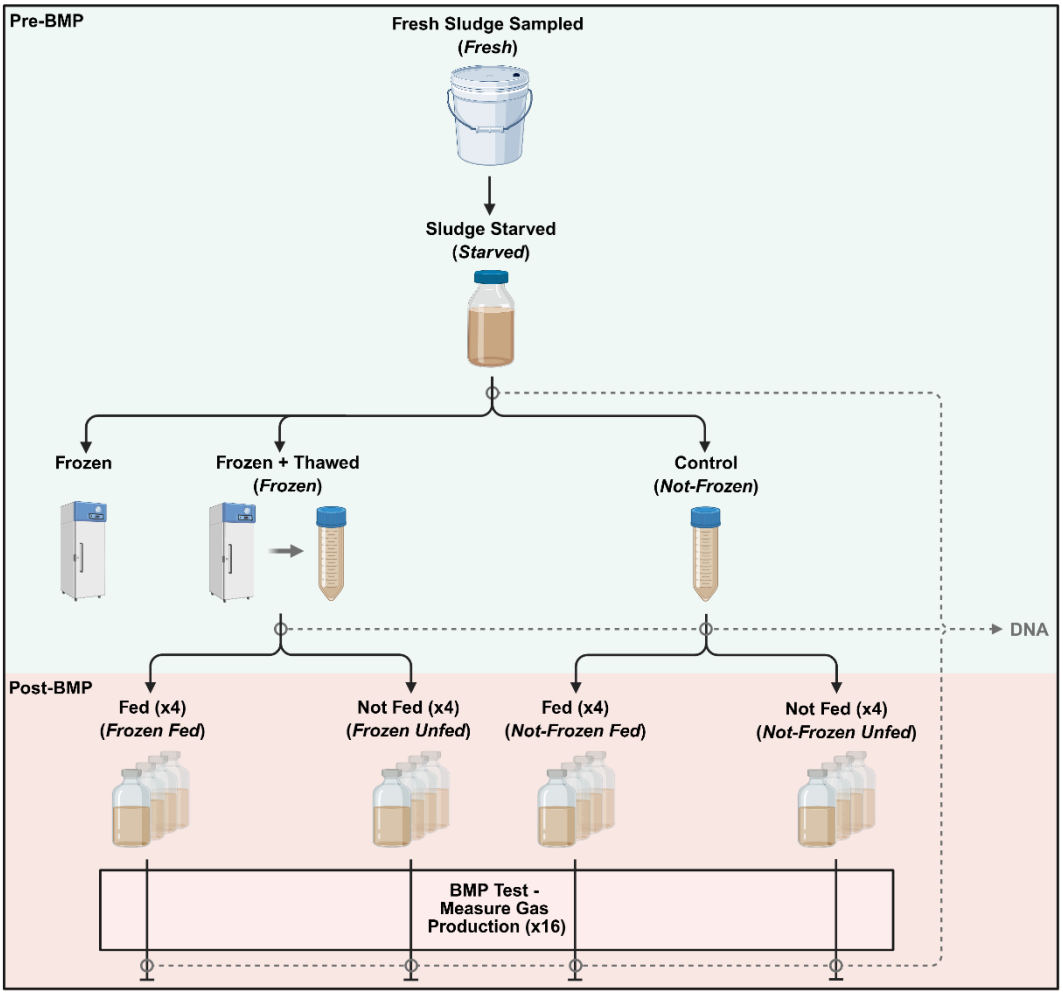


Figure 5.5: Sludge storage and viability testing process.
Overview of treatments and points of DNA sampling used for sludge viability testing and storage. Samples in the upper (green) panel are collectively referred to as pre-BMP test samples, those in the lower (pink) are post-BMP. Text in italics indicates naming of samples in the text. Points at which DNA was extracted and sequenced are indicated by circles. "Frozen + Thawed" and "Control" samples were taken forward to perform the BMP test, while "Frozen" samples (far left) were retained at -20 ^{o}C for further work. "Fed" refers to samples supplemented with cellulose for the BMP test.

5.2.2.1 Developing a reliable methane sensor

A reproducible method of methane quantification was required for the BMP test. An approach based on an infrared gas sensor was trialled and found to be able to reproducibly and rapidly measure the methane content of small (5 mL) gas samples (Figure A6).

5.2.2.2 Stored community viability test

Total gas and methane production were measured over a period of 52 days for the four quadruplicate BMP test populations (pink panel, Figure 5.5). Comparable amounts of total biogas (Figure 5.6, Panel A) and methane (Figure 5.6, Panel B) were produced in Frozen-Fed and Not-Frozen Fed samples by the end of the experiment, demonstrating that a viable community had survived storage via freezing.

While both fed treatments reached a similar end point in terms of total biogas (mean 501 mL +/- SE 11 mL for Frozen Fed and 496 mL +/- 14 mL for Not-Frozen Fed) and methane (mean 224 mL +/- 12 mL for Frozen Fed and 227 mL +/- 9 mL for Not-Frozen Fed) produced, there was a lag period from around 6 to 40 days where gas levels were lower in the previously frozen samples compared with their unfrozen counterparts, suggesting that the act of freezing had modified the initial performance of these communities. Unfed controls produced very little biogas (mean 38 mL +/- 3 mL for Frozen Unfed and 21 mL +/- 2 mL for Not-Frozen Unfed at 52 days) as had been expected given that the material used for these tests had been starved to deplete residual biodegradable material.

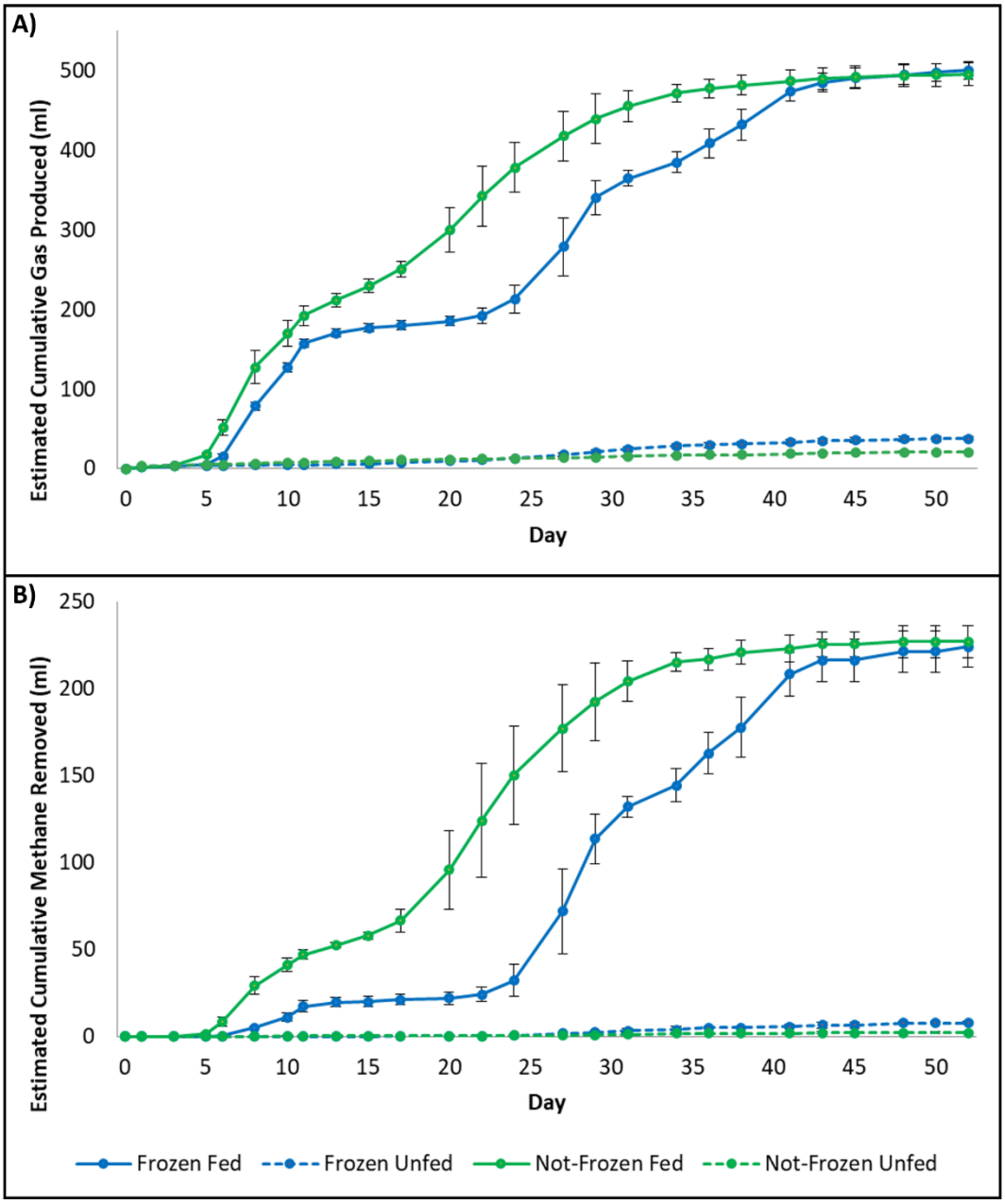


Figure 5.6: Gas production following sludge storage.
Cumulative gas production by storage method and feeding status in BMP test. **A)** Total biogas production **B)** methane production. Data points are means of a three to four (minimum three) biological replicates. Error bars are +/- one standard deviation. Blue lines = Frozen communities. Green lines = Not-frozen. Solid lines = Fed samples. Dashed lines = Unfed.

5.2.2.3 Community composition changes during storage

16S sequencing was carried out on DNA samples taken prior to and after the BMP test (Figure 5.5) to see if this could explain the differences observed in gas production dynamics between the different communities (Figure 5.6). Although a slightly higher

average number of ASVs were detected in Frozen Fed and Frozen Unfed communities (Figure 5.7), values were similar across all pre- and post- THP samples (Figure 5.7) indicating that the total number of different species was likely similar between all samples regardless of stage of the experiment or whether they had been frozen.

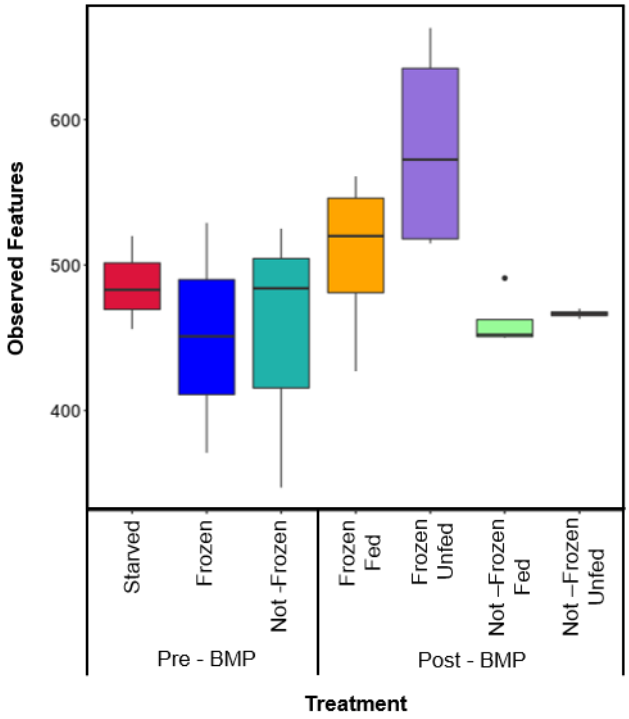


Figure 5.7: Observed features by sample.
Average number of observed features (ASVs) for the pre- and post-THP community treatments. Boxplots are medians of triplicate technical replicates (all pre-BMP treatments) or quadruplicate biological replicates (post-BMP). Boxes encompass quartiles 1 - 3. Whiskers are 1.5 x interquartile range. Points beyond whiskers are outliers. Adapted from figure produced by Novogene.

Comparison of the top 10 most abundant taxonomic groups at the genus level revealed relatively high levels of archaea compared to bacteria in all samples. Clear differences were present between the abundant post-BMP archaeal communities. In all except the Frozen Fed treatment, the archaeal community was composed of two dominant genera, *Methanosaeta* and *Methanospirillum*. The fact this was also the case in all samples taken prior to the BMP test suggests that this was the default state of the community. In all four replicate Frozen Fed communities a shift appeared to have taken place whereby

Methanosarcina had largely replaced the previously dominant archaeal groups. This was presumably triggered by the freezing, but feeding was also clearly required for this shift as Frozen Unfed communities remained in a state more similar to the other communities.

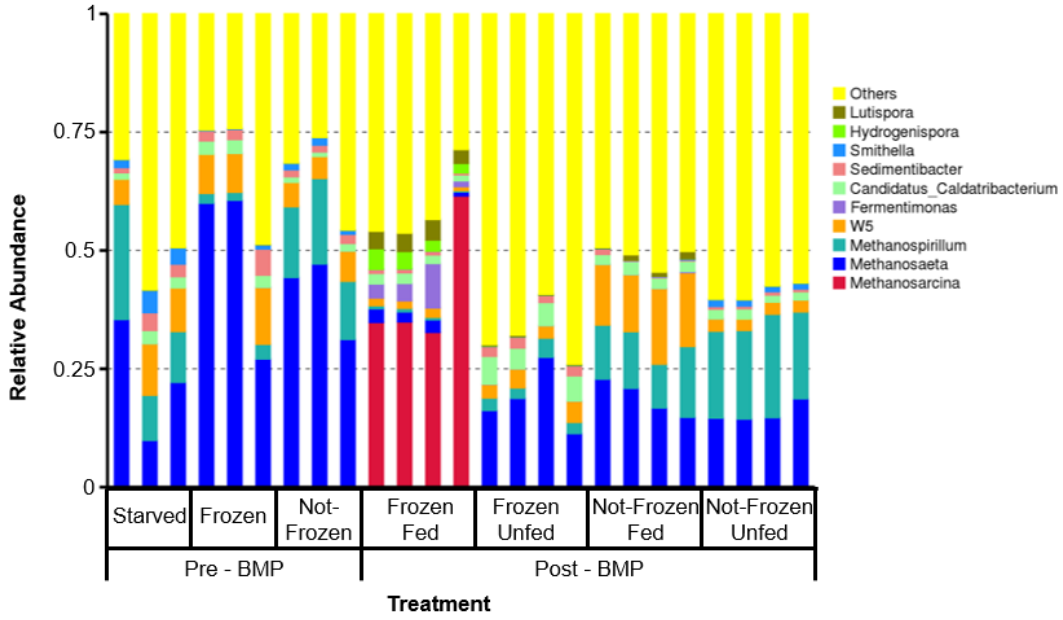


Figure 5.8: Taxonomic classification of abundant genera.
Top 10 most abundant genera across treatments at the genus taxonomic level. Predicted identity of each group is indicated to the right. Bar height is proportional to the relative abundance of reads derived from each genus from the sample. Bars are individual samples: technical replicates for all pre-BMP treatments and quadruplicate biological replicates for post-BMP treatments. Adapted from figure produced by Novogene.

Overall this demonstrated that a viable AD community could be recovered from frozen Esholt sludge relatively reproducibly. Additionally, stored sludge aliquots of this community were now available for further experiments which should serve as a more consistent starting point.

5.2.3 Colony picker pipeline development

One of the chapter objectives was to develop methods for isolating organisms in pure culture from AD sludge at higher throughput, as previous work (Chapter 4) had highlighted that the time taken to process culture plates by hand had been the primary limiting factor. Higher throughput methods using a colony picker were first tested on a laboratory strain of *E. coli* (MG1655) to ensure that they were capable of achieving sufficient contamination free growth before being used to isolate unknown organisms from sludge.

5.2.3.1 Isolation and liquid culture outgrowth

The colony picker was used to pick and array *E. coli* colonies from solid spread plates into 96 well plates containing liquid medium for outgrowth. An alternating pattern of wells was inoculated to detect possible cross contamination (Figure 5.9).

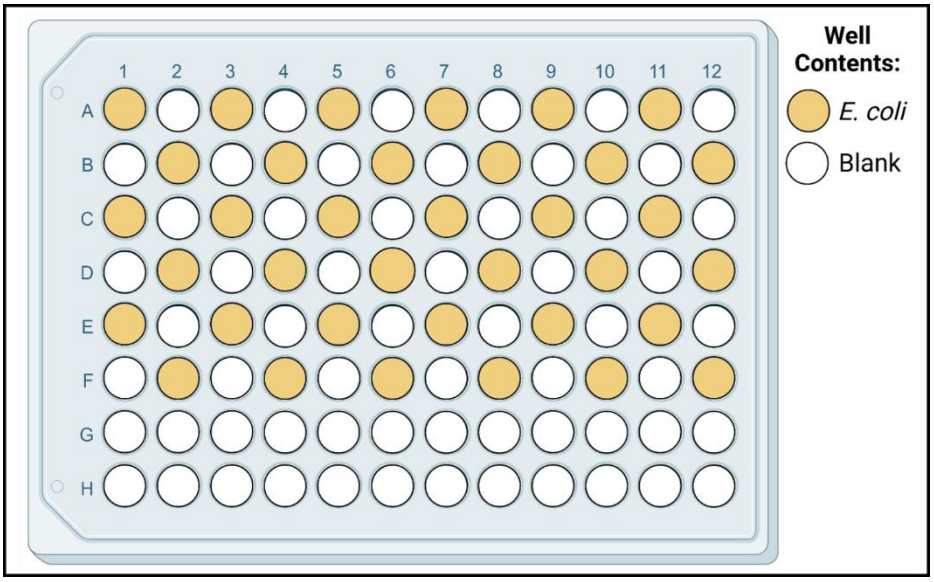


Figure 5.9: *E. coli* test array plate layout.
E. coli was inoculated into alternate wells in rows A-F starting with well A1 (orange circles). Blank (uninoculated) wells are alternate in rows A-F starting with well A2, and all wells in rows G and H.

Two different 96 well plate types (deep and standard (referred to here as "shallow" – see section 2.1.26, Methods) were inoculated in this way, sealed, and incubated using three different protocols (static, slow shaking and fast shaking (see section 2.1.26.2 Methods) (Figure 5.10).

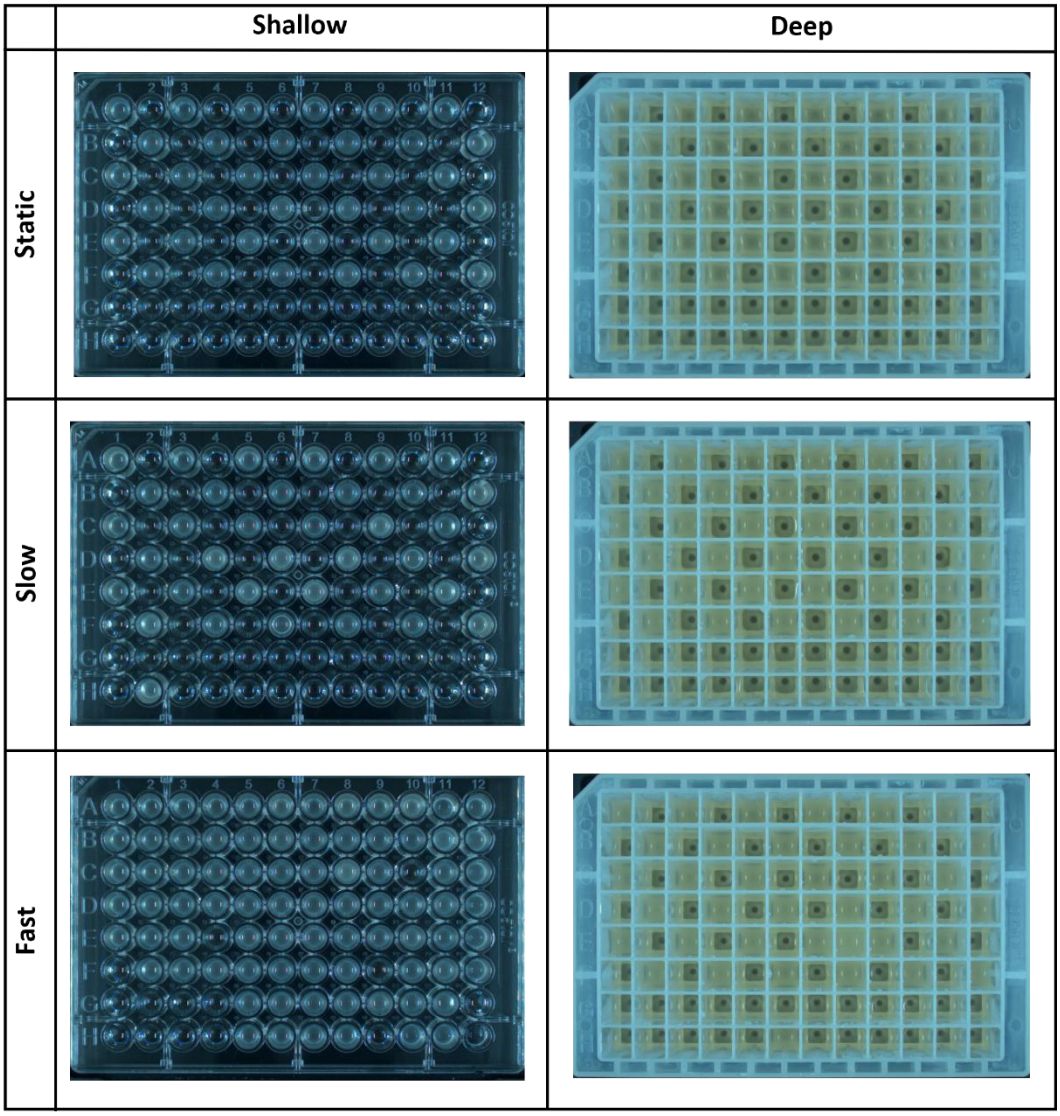


Figure 5.10: *E. coli* liquid culture outgrowth.
Results of *E. coli* inoculation and outgrowth in two different plate types incubated at three different speeds. Plates were initially arrayed as explained in Figure 5.9. Turbid wells indicate locations at which *E. coli* was able to grow.

Following overnight incubation contamination was detected in both plate types incubated with fast shaking, and the shallow plate incubated with slow shaking. Neither plate contained any contaminated wells when incubated static, and the deep well plate which had been incubated with slow shaking was also free of contamination. Inoculated wells in statically incubated cultures were visibly less turbid. This data suggests that contamination was a result of overspill between wells through the plate seal as it was worsened by shallower plates and faster incubation speeds. The conclusion for the protocol was that static incubation should be used wherever possible, but if turbid culture

is required (for example for DNA extraction) slow shaking in deep well plates should be used to avoid contamination.

5.2.3.2 Storage and revival

A glycerol cryostorage and revival method suitable for handling multi well plates was next developed. In section 5.2.3.1 statically incubated shallow plates had been free from contamination, and given that these would take up less space in the -80 ^{o}C freezer these were used as the basis for initial testing. Three replica shallow 96 well plates were arrayed as shown in Figure 5.9 and frozen. These were revived into triplicate shallow plates using three different methods: direct transfer of still frozen material, or by defrosting followed by liquid transfer by a pipette or the colony picker (see section 2.1.26.3, Methods). Outgrowth following revival revealed an unsatisfactory level of contamination for all three methods (data not shown).

The test was repeated using deep well plates. This time only revival via defrosting and rearraying using the colony picker was tested as it had been found to be substantially faster than the other methods. All three replica deep well plates generated using this method had growth in all expected wells and were free from contamination following incubation. This outcome suggested that deep well plates should be used for all parts of this protocol including storage as they appeared to have been less susceptible to contamination.

5.2.4 Anaerobic plates

In an attempt to find a way to combine the high throughput of the colony picker with the ability to isolate anaerobic organisms outside of the anaerobic chamber, an approach for plating anaerobic organisms in aerobic conditions which had been reported in the literature was investigated.

5.2.4.1 Separate autoclaving with antioxidants

A protocol based on solid medium with separately autoclaved components additionally supplemented with antioxidants was trialled as it had been reported to support a wide variety of anaerobic organisms including obligate anaerobes (Dione *et al.* 2016). This method used two plate types, R medium which contains the solid medium components

autoclaved separately plus a mixture of antioxidants, and Schaedler agar (SA), which is the base components of R medium autoclaved together with no supplementary components. The expectation was that if R medium supported growth of both anaerobes and aerobes many more colonies would be recovered on sludge spread on these plates than for the control medium (SA). This difference should be detectable by plate counting.

During R medium preparation the recipe had to be adjusted as the haemin component of the antioxidant solution was found to dissolve poorly and block the filter membrane, and was instead prepared separately by autoclaving. Fresh sludge was spread on both plate types. Unexpectedly, no colonies were recovered on R medium plates after four days of incubation (Figure 5.11, Panel B), whereas control SA plates (expected to support fewer organisms) had vigorous growth (Figure 5.11, Panel A). Even after 17 further days of incubation no colony was detected on any R medium plate.

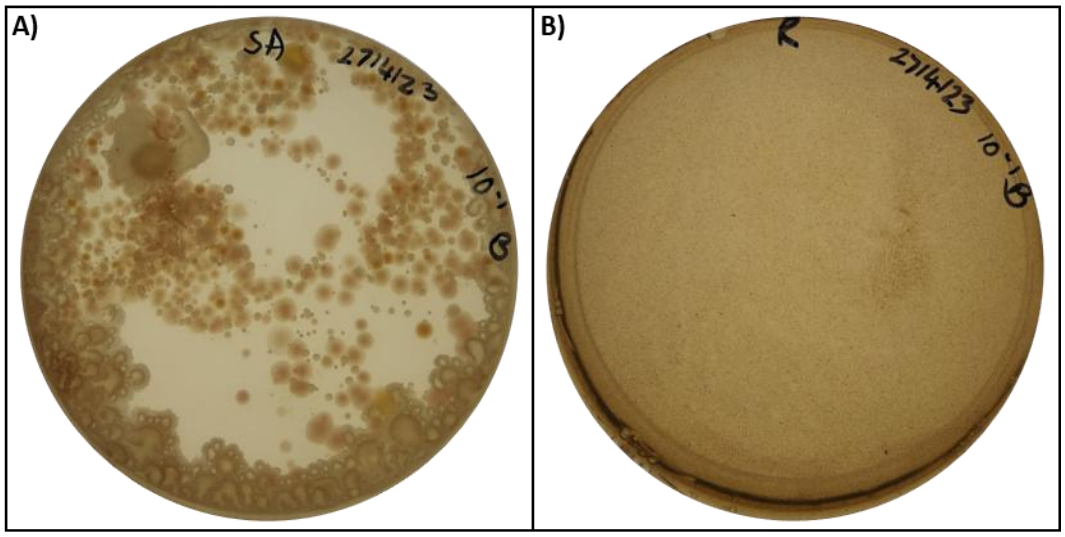


Figure 5.11: Colony recovery of sludge on supplemented and control plates. Representative images of Esholt sludge plated on the two different plate types. **A)** Schaedler Agar (control). **B)** R medium.

It appeared that some factor present in the R but not the control medium was lethal to microorganisms from the sludge. The possibility that some component had been included at the wrong concentration was ruled out by careful rechecking of the protocols.

Given that issues had been encountered with the haemin component of the R medium, a second experiment was carried out in which the concentration of this was reduced to that

present in normal Schaedler's agar. All other components were kept the same. Once again, no growth was recovered on any R medium plate while vigorous growth was present on SA, indicating that some other undetermined factor was responsible.

5.2.4.2 Separate autoclaving only

As efforts involving supplementary addition of antioxidants had yielded unexpected results, a simpler method based only on autoclaving media components separately was attempted. This method used two media which each contain the same components, but in one the agar is autoclaved separately to phosphate (S medium) and in the control all components are autoclaved together (T medium) (Kawasaki and Kamagata 2017).

The expectation was that S medium plates would support more organisms from sludge and therefore yield higher colony counts than T medium when incubated aerobically. To test whether transient oxygen exposure followed by anaerobic incubation had any impact on colony recovery, an additional set of each type of plate was prepared aerobically but then transferred to the anaerobic chamber for incubation.

Periodic plate inspection revealed fewer and smaller colonies than had been expected in all conditions, so the incubation period was increased to seven days, before colonies were counted. Aerobically incubated plates had larger more defined colonies which were generally clearly visible on a light box (Figure 5.12, Panels A and B). Significantly more colonies were recovered on S plates incubated in this condition (mean 12.6 x 10^3 CFU/mL +/- SE 0.617 x 10^3) than T (6.23 x 10^3 +/- SE 0.273 x 10^3) (t-test: t = 9.48; d.f. = 4; p = 0.0006902) (Figure 5.12, Panel C).

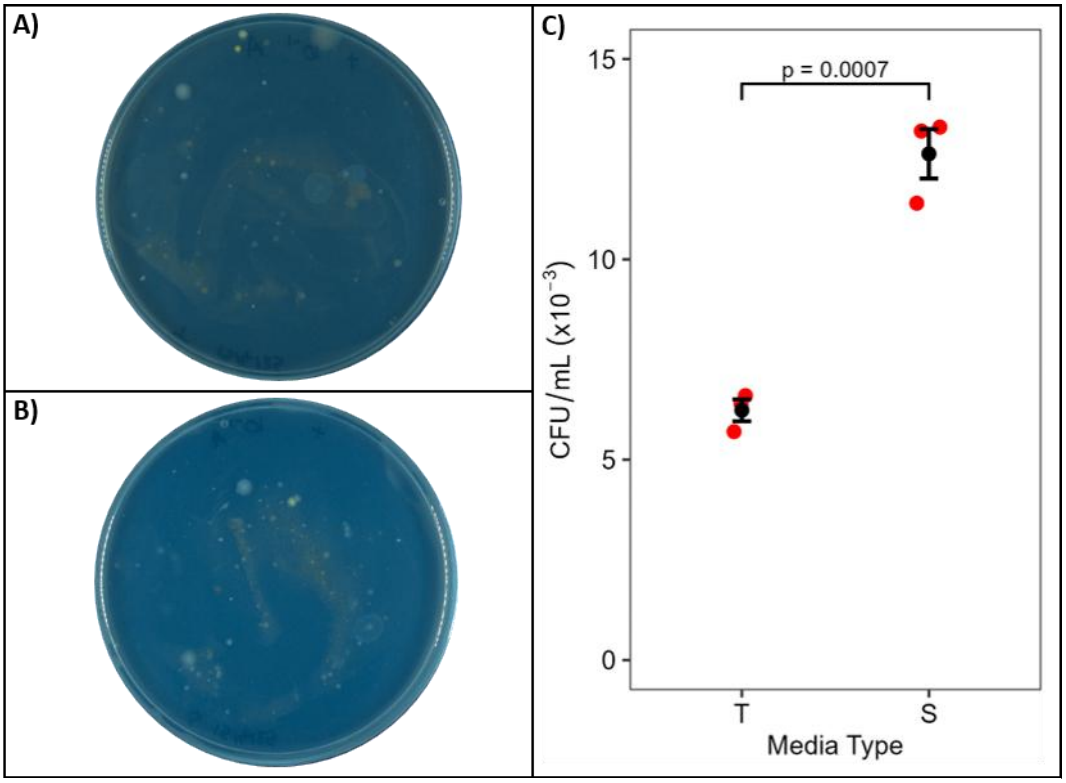


Figure 5.12: T and S medium comparison (aerobic incubation).
A) Representative image of colonies on aerobically incubated T plate. **B)** Representative image of colonies on aerobically incubated S plate. **C)** Mean colony forming units per mL of wet sludge (black points) calculated from triplicate plate counts (red points) after seven days of aerobic incubation. Mean CFU/mL were significantly higher on S medium (12.6 x 10^3 +/- 0.617 x 10^3) compared with T (6.23 x 10^3 +/- 0.273 x 10^3) medium; t = 9.48; d.f. = 4; p = 0.0006902. Error bars are +/- one standard error.

Colonies were generally smaller and more difficult to distinguish on anaerobically incubated plates (Figure 5.13, Panels A and B). Those that could be seen under a light box were counted and found to not significantly differ in number between plate types (Figure 5.13, Panel C). There were many small translucent colonies which were not possible to accurately count on the anaerobic plates, which means that direct CFU comparisons between the two incubation conditions could not be made reliably.

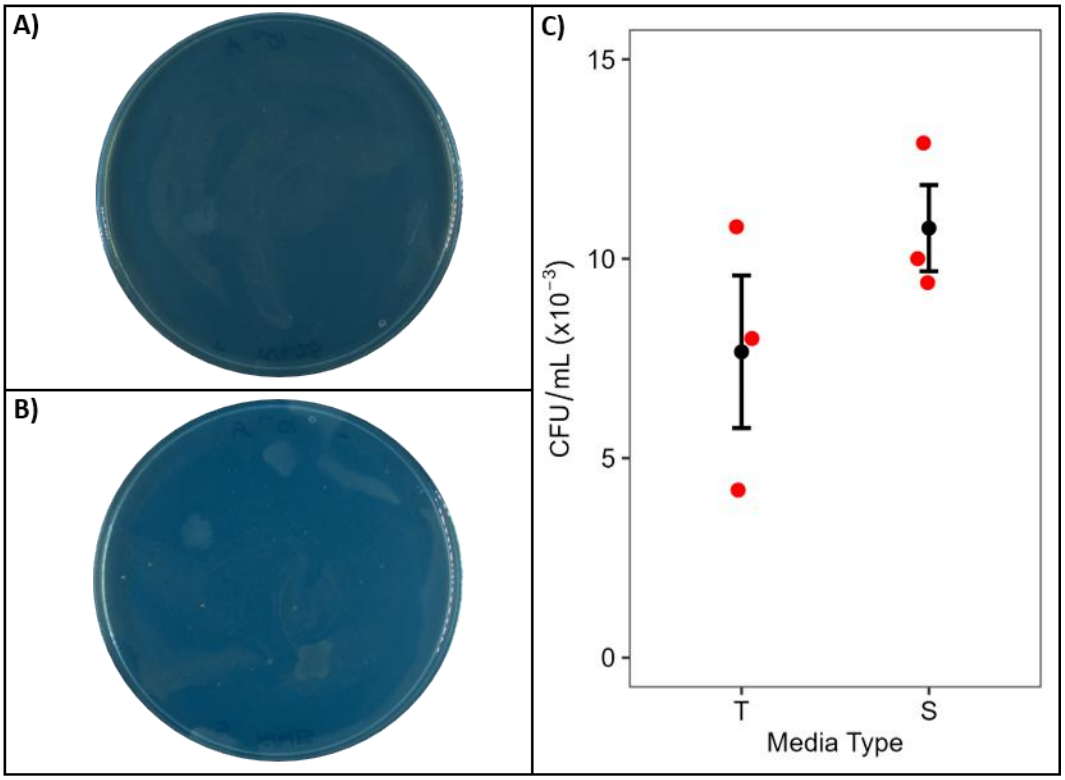


Figure 5.13: T and S medium comparison (anaerobic incubation).
A) Representative image of colonies on anaerobically incubated T plate. **B)** Representative image of colonies on anaerobically incubated S plate. **C)** Mean colony forming units per mL of wet sludge (black points) calculated from triplicate plate counts (red points) after seven days of anaerobic incubation. Mean colony counts were not significantly between T and S medium; t = 1.411; d.f. = 4; p = 0.231. Error bars are +/- one standard error.

Small irregular colonies (Figure A7, Appendix) were present on the majority of plates across all conditions. They were more numerous on plates spread with more dilute sludge and were absent from the least dilute plates, which were used to generate the CFU figures and for plate washing (see below). The colonies were present but less numerous on control plates, with one exception which had been accidentally opened in the air lock of the anaerobic chamber early in the experiment. Based on this evidence these colonies may have been contaminant airborne bacteria which were able to opportunistically colonise plates with few competitor organisms on during periods of manipulation (such as imaging).

The plating work suggested that S medium was able to support a significantly greater number of organisms than T plates in aerobic but not anaerobic incubation conditions,

which is what would have been expected if this method of preparing the medium allowed a proportion of the anaerobes present in the sludge to grow.

To investigate this further, platewashes of S and T medium plates incubated in both aerobic and anaerobic conditions were DNA extracted and analysed by 16S sequencing to determine the identity of the colonies that had been present on them. Platewashes were also attempted for the control plates to determine the identity of the contaminant colonies, but insufficient DNA was recovered from these for 16S sequencing. A control DNA sample which had been extracted from the sludge used to generate these plates was included as a reference for the original community.

Alpha diversity metrics showed sludge had the highest Shannon index and total number of ASVs (Table 5.2) as expected as this sample would have contained DNA from the whole community including those organisms not recoverable on plates. Shannon index across all platewash samples was similar. The lowest diversity in terms of ASVs was found in aerobically incubated T medium, intermediate on S plates incubated aerobically and highest in both plate types incubated anaerobically. In agreement with the counting data these results were as expected if the separately autoclaved (S) medium was able to permit some anaerobes to grow in aerobic incubation conditions.

Table 5.2: Alpha diversity metrics for plate was 16S data.
Values are derived from 16S sequencing data from a single DNA sample from each condition. Sludge is derived from a DNA sample extracted from the material used to generate the T and S plates.

Condition	ASVs	Shannon Index	Pielou_e
T Aerobic	78	4.63	0.74
S Aerobic	136	4.32	0.61
T Anaerobic	160	4.21	0.58
S Anaerobic	157	4.22	0.58
Sludge	709	6.84	0.72

Cross comparison of shared ASVs between different conditions (Figure 5.14, Panel A) revealed a general trend of higher levels of cooccurrence based on incubation method (i.e. aerobic/anaerobic) compared with media type. In agreement with this, analysis of the most abundant groups at the genus level (Figure 5.14, Panel B) showed the

composition of the most abundant community members appeared to mainly be dictated by incubation condition and varied little between plate types in the same condition. The composition of anaerobically incubated plates appeared more similar to that of sludge than aerobically incubated plates, indicating that transient anaerobic incubation was the more important factor in terms of permitting the growth of organisms more representative of the original community.

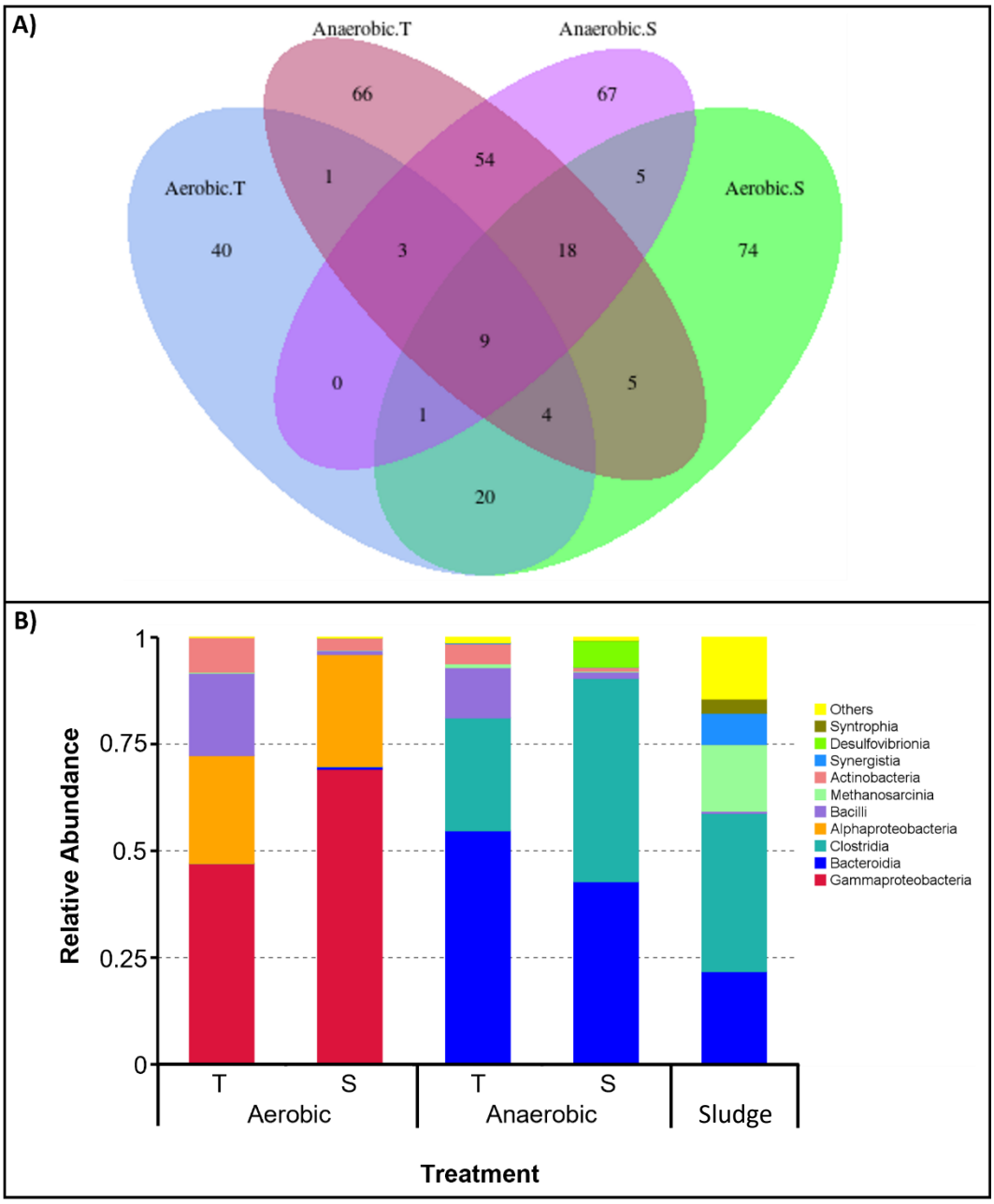


Figure 5.14: T and S medium plate wash 16S analysis.

A) ASVs shared between or unique to S and T plates incubated aerobically and anaerobically. Numbers indicate ASVs, either shared between treatments (overlapping sectors) or unique to them (outside sectors). **B)** Top 10 most abundant groups across treatments at the class taxonomic level. Sludge is derived from a DNA sample extracted from the material used to generate the T and S plates. Predicted identity of each group is indicated to the right. Bar height is proportional to the relative abundance of reads derived from each genus from the sample. Groups are derived from 16S sequencing data from a single sample. Figures produced by or adapted from Novogene.

Collectively these results suggested that while the separately autoclaved media may have supported plate growth of a broader proportion of the community in aerobic conditions, the best results in terms of total diversity and proportional representation of abundant groups was achieved by incubating the plates in anaerobic conditions regardless of the method used to prepare the media.

5.2.5 High throughput colony isolation from stored sludge

The colony picker pipeline (section 5.2.3) was next tested on the stored Esholt sludge (section 5.2.2). As the anaerobic plates tested (section 5.2.4) had not performed as expected, for this initial run plating was done in aerobic conditions using three media which had been used previously in Chapter 4 (nutrient agar (NA), brain heart infusion agar (BHI) and reinforced clostridial agar (RCA)) plus an additional one (anaerobic basal agar (ABA)).

Previous work with fresh sludge had shown that an ideal length of incubation on three of these media was four days (Chapter 4), but inspection of stored sludge spread plates after two days showed colonies had unexpected growth characteristics. There were in general fewer colonies than would be expected but these were large and on some of the lower dilution plates had already begun to merge, while more dilute plates had few or no colonies on. The plates were incubated for a further five days to see if slower growing colonies would emerge on the more dilute plates, but this did not happen and instead those already present continued to grow. As it was not possible to accurately count or pick these colonies these plates were discarded.

The experiment was repeated but this time plates were imaged and counted after one day of incubation to avoid overgrowth (Figure 5.15, Panel A). There were significantly more colonies in ABA plates than BHI and RCA, but in general there were no dramatic differences between plate types (Figure 5.15, Panel B). Given the lower total number of colonies, colony counts were only above 30 (usually the lower limit for plate counting) on two of the 12 most concentrated plates (both ABA), so these comparisons should be viewed with caution.

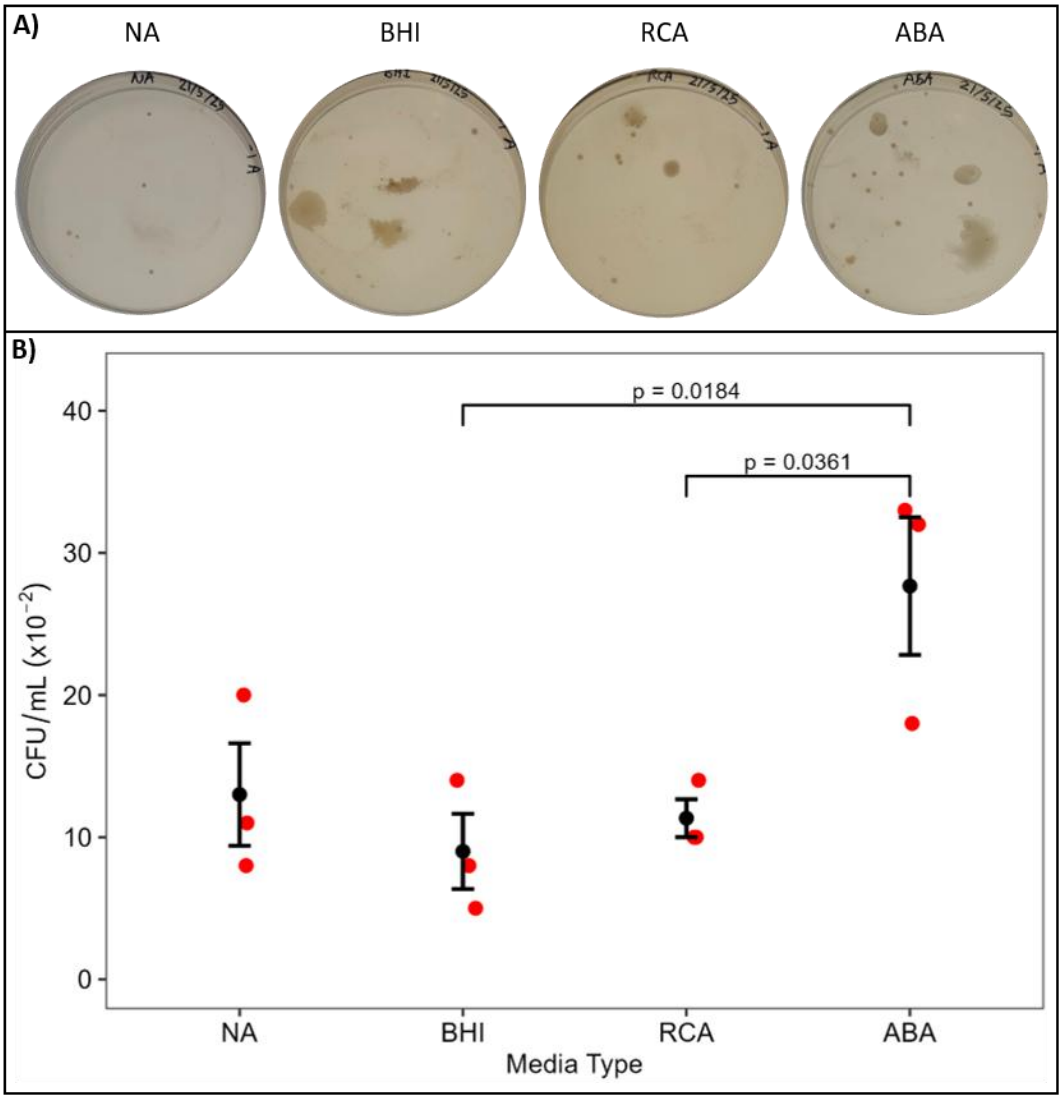


Figure 5.15: Plated stored sludge colony morphologies and counts between media. **A)** Representative images of colonies on the four different media used. **B)** Mean colony forming units per mL of wet sludge (black points) calculated from triplicate plate counts (red points) after one day of aerobic incubation. There was a significant difference in CFU between media types ($F = 6.299$; d.f. = 3, 8; $p = 0.0168$). Post-hoc testing showed there were significantly higher CFU values for ABA (mean 27.7 CFU +/- se 4.84) compared with RCA (11.3 +/- 1.33 CFU; $p = 0.0361$) and BHI (9 +/- 2.65 CFU; $p = 0.0184$). No other pairwise differences were statistically significant. Error bars are +/- one standard error. NA = nutrient agar. BHI = brain heart infusion agar. RCA = reinforced clostridial agar. ABA = anaerobic basal agar.

Colonies were picked into each of their respective liquid media types in separate deep well plates using the colony picker. All plate dilutions were used for this to provide as many colonies as possible (Table 5.3). Each plate contained six control wells to detect

contamination: three of which had been inoculated from a control plate containing no organisms and three that were not inoculated at all.

Plates were incubated statically for four days, and colony growth in wells was determined by visible inspection. Recovery was generally high, with the lowest level for NB at 84% and highest for ABB at 97%. These figures are also likely an underestimate as it was previously observed that some cultures without visible growth did contain viable cells (Chapter 4).

None of the dedicated control wells in each plate contained any growth. Contamination was detected in a single empty well in the BHI plate. If these empty wells are considered as additional uninoculated control wells then this contamination only occurred in one well of 92 across all plate types (i.e. around a 1% contamination rate) which is tolerable for this purpose.

Table 5.3: High throughput liquid culture recovery of picked colonies.
Total number of colonies picked into 96 well plates for each medium type. "Wells with visible growth" is the number of wells inoculated where growth was visually detectable after four days of static incubation. Recovery is the percentage of inoculated wells of each medium type with visible growth. NB = nutrient broth. RCM = reinforced clostridial medium. BHI = brain heart infusion broth. ABB = anaerobic basal broth.

Medium	Wells Inoculated with Colonies	Wells with Visible Growth	Recovery (%)
NB	80	67	84
RCM	53	46	87
BHI	57	52	91
ABB	90	87	97

These plates were frozen at -80 ^{o}C for future studies.

5.3 Discussion

5.3.1 Aims Recap

The aims of this chapter were to determine whether the sludge used throughout this project could be stored by freezing while maintaining a functional AD community, and to test methods for higher throughput isolation and culturing of AD microorganisms using an automated colony picker.

5.3.2 Chapter achievements

Work on characterising the sludge used throughout this project generated unexpected results which further justified efforts to preserve a digester community. It was possible to freeze and revive a functional AD community from Esholt sludge as demonstrated by methane production in all previously frozen fed replicate communities assayed by BMP testing (Figure 5.6). Additional aliquots of this sludge remain in storage and are available for use in further experiments. Previous "by hand" methods for isolating, culturing and storing organisms from AD sludge were also successfully adapted for use with a colony picker, vastly increasing potential throughput. Efforts to use previously described methods for growing anaerobic organisms on plates in aerobic conditions had some unexpected outcomes.

5.3.3 System feed investigations

The expectation had been that post-THP feed would contain fewer or no organisms compared with pre-THP due to cell death in the pretreatment process. This was apparently not the case as total numbers recovered on plates were actually similar (Figure 5.1). It has been reported that THP pretreatment sterilises sludge (Barber 2016; Svennevik *et al.* 2019; Haffiez *et al.* 2022) which is defined as the complete destruction of all forms of microorganisms (Rutala and Weber 2015). However, direct evidence of this is surprisingly difficult to find, possibly because the majority of THP literature focuses on the chemical impacts of the treatment on feed and its subsequent digestibility (Carrère *et al.* 2010; Pilli *et al.* 2015; Barber 2016) while studies addressing biological impacts often use DNA based molecular methods (Berde and Jolis 2018; Tong *et al.* 2019; Zhang *et al.* 2021) which do not distinguish between live and dead cells. Studies specifically measuring microbial viability after THP pretreatment have reported reductions in *E. coli*, coliforms

and other Gram-negative bacteria (Redhead *et al.* 2020) and *E. coli* and *C. perfringens* (Levantesi *et al.* 2015) to undetectable levels, which differ from the results obtained here.

Further investigations using 16S sequencing suggested the post-THP feed investigated in this project was instead markedly less diverse. The simplest explanation for this is that only a subset of the organisms with mechanisms suitable for survival in extreme conditions such as spore formers (Mirsoleimani Azizi *et al.* 2024) were able to survive thermal hydrolysis and make it through. If this were the case, some explanation needs to be provided as to why the most numerous organisms detected by post-THP DNA analysis are not known to be spore formers (see section 5.2.1.2). One possibility is that these community members could not be recovered in the plating conditions used, and the organisms that were recovered on the plates (and were therefore responsible for the counts) were spore formers that were actually comparatively rare in the post-THP community (part of the "other" group in the DNA data (Figure 5.4)). Attempts to directly sequence DNA just from the subset of organisms grown on the plates to verify this were unsuccessful, possibly due to the requirement of an additional PCR amplification to process this very low yield DNA. A second possible explanation is that the organisms which grew on the plates were able to survive the harsh pretreatment conditions through another mechanism besides spore formation. For example, some bacteria may have been protected from the harsh conditions by being embedded in solid matrices (Mackey *et al.* 2006; Otte *et al.* 2018) such as particles or flocs in the feed.

Prior to these experiments, the assumption that THP feed was sterile had informed the decision to use THP fed digesters as the basis for this project, using the reasoning that this cuts out a potential source of ingress of organisms (Campanaro *et al.* 2020), therefore leading to potentially less variable communities in terms of species composition (compared with digesters fed non-pretreated feed). The data generated here suggest that while the feed was not sterile, it was less diverse at the species level. It is possible that the digester community is less varied as a result of this, however further experiments involving sampling of the digester itself across multiple different timepoints, and comparison to a suitable control digester with non-pretreated feed would be required to determine this. Overall, this work revealed that the digester community sampled throughout this project may be less simple than originally thought, which further justified

the value of investigating freezing as a means of storing a mixed microbial community from an anaerobic digester.

5.3.4 Sludge storage community changes

After a lag period, AD communities from sludge which had been frozen and revived were able to produce comparable amounts of methane to their non-frozen counterparts (Figure 5.6). Monitoring of these communities by 16S sequencing revealed substantial shifts in their archaeal members, with *Methanosarcina* almost entirely replacing both previously dominant *Methanosaeta* and *Methanospirillum* in this condition only (Figure 5.8).

Competition between *Methanosarcina* and *Methanosaeta*, which are both acetoclastic methanogens (Dyksma *et al.* 2020), has been reported previously (Conklin *et al.* 2006; Bonk *et al.* 2018). *Methanosaeta* has a higher affinity for acetate than *Methanosarcina* but a lower growth rate, so tends to dominate when acetate levels are low and conditions are stable (Conklin *et al.* 2006; De Vrieze *et al.* 2012). *Methanosarcina* is unusual among methanogens in that it can grow in clusters and has a comparatively larger cell size, making it more resilient to challenging conditions (De Vrieze *et al.* 2012). These differences provide a possible explanation for the replacement of *Methanosaeta* with *Methanosarcina* only in Frozen Fed communities. If a higher proportion of *Methanosarcina* were able to survive freezing than *Methanosaeta* on account of their increased robustness, this would have put the two competitors on a more even footing when resources became plentiful on feeding after thawing, allowing *Methanosarcina* to outcompete its rival *Methanosaeta*. In the Non-Frozen Fed communities *Methanosaeta* would have remained dominant such that even when resources became plentiful (which would normally favour *Methanosarcina* due to their increased growth rate (De Vrieze *et al.* 2012; Yin *et al.* 2022)), this shift did not occur.

Methanosarcina appears to have additionally replaced the dominant hydrogenotrophic methanogen *Methanospirillum* (Humm *et al.* 2025) as well as the acetoclastic *Methanosaeta* in Frozen Fed communities. *Methanosarcina* is the only methanogen known to be capable of both types of methanogenesis (Galagan *et al.* 2002). One possibility is that it assumed both roles after outcompeting the two original niche holders

in these communities, although this has not to the author's knowledge been reported before.

A possible concern is whether the frozen samples prepared here would remain viable over a longer period of time, since these tests were conducted soon after freezing. While changes in stored AD community composition and/or performance over time have been noted by others (Hagen *et al.* 2015; Astals *et al.* 2020), in both cases if a viable community could be resuscitated after a short time period (days) a viable (albeit less productive) community was also recoverable after longer time periods of weeks (Astals *et al.* 2020) or months (Hagen *et al.* 2015). Further work could investigate the impacts of longer storage periods for the sludge used in this project.

5.3.5 Anaerobic plates

Plates containing separately autoclaved components supplemented with antioxidants (R medium, (Dione *et al.* 2016)) intended to permit growth of anaerobes in the presence of oxygen were lethal to microorganisms in this work (Figure 5.11). The causative factor must be something that was present in the R medium but not the control (SA) since organisms grew well on these plates. Troubleshooting ruled out haemin as the culprit, which means one of the other components of the antioxidant solution of the R medium was likely the cause. One possible explanation is an issue with one of the chemicals itself, e.g. from a bad batch, or contaminated with some other toxic component. One way to determine the cause would have been to produce and test multiple drop out media containing all but one of the component chemicals. As these plates were time consuming and labour intensive to produce, and alternative protocols were available, this was deemed an uneconomical use of time and not taken further.

Results were closer to expectations when a different protocol with no additional antioxidants was used (Kawasaki and Kamagata 2017). The data suggested that the separately autoclaved version of this medium (S) was able to support plate growth of a larger number (Figure 5.12) and diversity (Table 5.2) of organisms than the control (T) medium when incubated aerobically. However, even better results in terms of diversity of organisms recovered (Table 5.2) and closeness of match to the progenitor community (Figure 5.14) were achieved by incubating these plates anaerobically regardless of how

the medium was prepared. Transient anaerobic incubation is faster and simpler than separately preparing the medium if anaerobic incubators are available. Ultimately, the best solution which guarantees suitable conditions for even strict anaerobes is to perform all parts of the process in an anaerobic chamber, which would be the recommendation for a high throughput campaign.

5.3.6 Colony picker protocol

Low throughput manual methods for colony isolation and storage from sludge used elsewhere in this project (chapter 4) were adapted into a smaller scale and higher throughput protocol based on multi well plates using a colony picker. Initial refinement of this protocol using *E. coli* flagged up major contamination issues that would have been highly compromising to actual AD organism isolation experiments. Preliminary testing of this protocol on sludge samples revealed it was substantially faster, requiring a fraction of the hands on time that had been required when performed manually (Chapter 4). Due to time limitations it was not possible to test the revival from glycerol stock portion of this protocol on unknown organisms. This would constitute immediate further work.

Interestingly the stored material (i.e. frozen aliquots prepared in section 5.2) appeared to yield fewer, generally more vigorously growing colonies on plates compared with fresh sludge. One possible explanation is that the act of freezing has selected for more robust community members while less resilient members had been lost. While this would be accompanied by some loss of diversity if this were the case, this could actually be advantageous in terms of recreating a defined AD community as resilient organisms able to survive storage and other harsh conditions may be more desirable in a SynCom.

5.3.7 Further work: generating and testing an AD library

With a high throughput protocol for AD organism isolation now in place, further work in this area would involve building up a larger library of stored organisms through additional rounds of plating and isolation, likely including a wider range of media types and incubation conditions to encourage a broader diversity of organisms to grow for isolation.

The isolation process could be combined into a single pipeline with library testing (Figure 5.16). Test AD SynComs (i.e. collections of isolates from the library) could be pooled in

combinations, inoculated into sterile sludge and tested for their ability to produce methane through BMP testing. Any such community found to be able to do so would have all component species required for AD and would provide a basis for further investigation. If methane were not produced, the pipeline could be iterated with further rounds of organism isolation, library growth and testing (Figure 5.16) with the aim of filling the missing niches in the synthetic community.

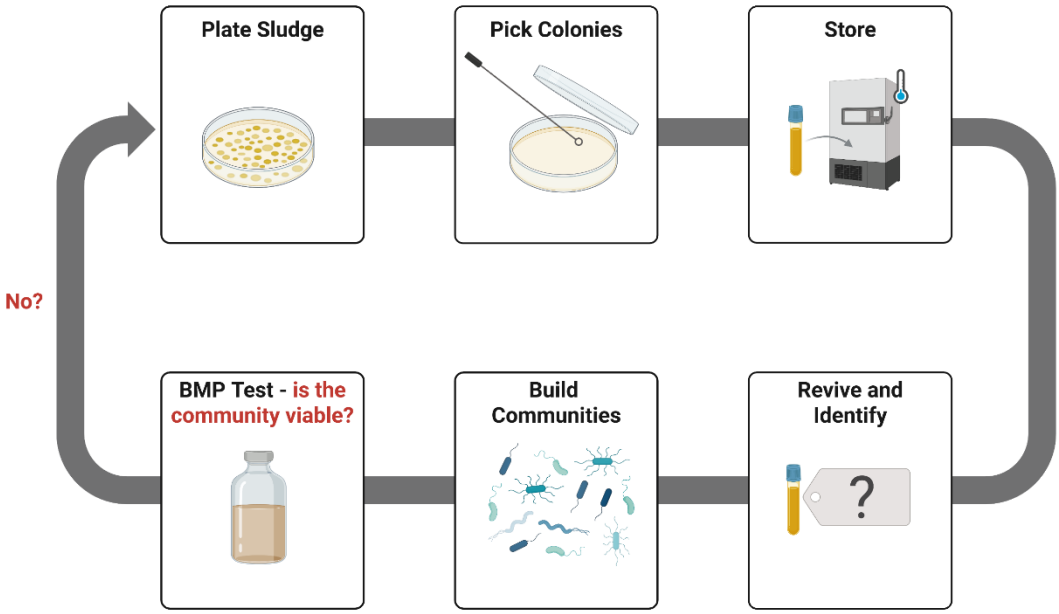


Figure 5.16: Proposed approach for generation AD SynComs.
Different plating techniques are used to obtain and store isolates of distinct AD organisms, which are revived and identified. Once a large enough collection of organisms is obtained, communities are built by mixing large numbers of revived microorganisms. Community viability is determined by methane production through BMP testing. If the community is not viable, further plating of sludge is conducted to isolate additional species and the cycle is repeated.

An anticipated issue is that some essential AD community members such as methanogens may be challenging to isolate using random plating. This issue could be addressed by a more targeted phase of library "gap filling", for example by using selective enrichment methods which have been described for methanogens (Duller *et al.* 2024). Another possibility could be to use a challenge and engraftment method recently described (Cheng *et al.* 2022). This would involve challenging an incomplete (i.e. non-methane producing) synthetic AD community of known composition with a wild AD community followed by

sequencing. Empty niches would be expected to be filled by donor community members which would simultaneously identify and fill gaps in the initial synthetic community.

Chapter 6: General Discussion

6.1 Aims recap

The overall aim of this project was to investigate biological approaches for monitoring and manipulating microbial communities in AD. This was explored in a number of different ways. In Chapter 3 an attempt was made to develop a whole cell volatile based bioreporter system which would be suitable for AD. In Chapter 4 methods for obtaining or generating more robust AD chassis organisms were investigated. In Chapter 5 the use of culturomics based approaches were explored as a first step towards developing an AD SynCom.

6.2 Achievements and challenges

Significant steps towards these goals were achieved in several areas. Directed evolution to increase the fitness of a laboratory microorganism in a parameter relevant to AD (i.e. the anaerobic environment) was demonstrated. Selective plating approaches were used to successfully obtain at least one close relative of a model microorganism which would represent a good candidate for the attempted transfer of genetic tools. Progress towards achieving AD culturomics was made, including the demonstration of a high throughput isolation and culturing protocol for AD organisms and the development of an identification approach based on full genome sequencing of pooled samples.

Some areas were found to be more challenging to progress, such as the development of volatile based reporters. However, refocusing of this portion of the project towards exogenous anaerobic ethylene synthesis generated some potentially encouraging data. Additional interesting findings from this work include data that suggested pre- and post-THP feed contained similar numbers of organisms but differing levels of community diversity. Investigation of frozen sludge communities yielded evidence of possible functional replacement of a hydrogenotrophic and acetoclastic methanogen pair with a single organism capable of both.

6.3 The potential of combined approaches

While different biological approaches for monitoring and manipulating microbial communities in AD were investigated separately throughout this project, with further

development many of them could theoretically be used in combination. For example, an endogenous organism with desirable characteristics isolated from AD could be further augmented, perhaps through directed evolution to increase performance with respect to some desirable parameter, or through engineering, for example by imparting the ability to utilise an unusual metabolic substrate (Figure 6.1). The organism could be reintroduced to a digester alone or as part of a mixture of other similar isolates. These organisms, which would already be expected to be well suited to this environment, would have the additional advantage of being able to obtain energy from an exogenously supplied substrate not usable by other community members, improving their stability (Shepherd *et al.* 2018) in the digester environment. A desirable mixture of organisms could be stabilised through engineered interdependence, for example by using engineered auxotrophy (Mee *et al.* 2014; Noto Guillen *et al.* 2021) such that different organisms need each other to be able to grow.

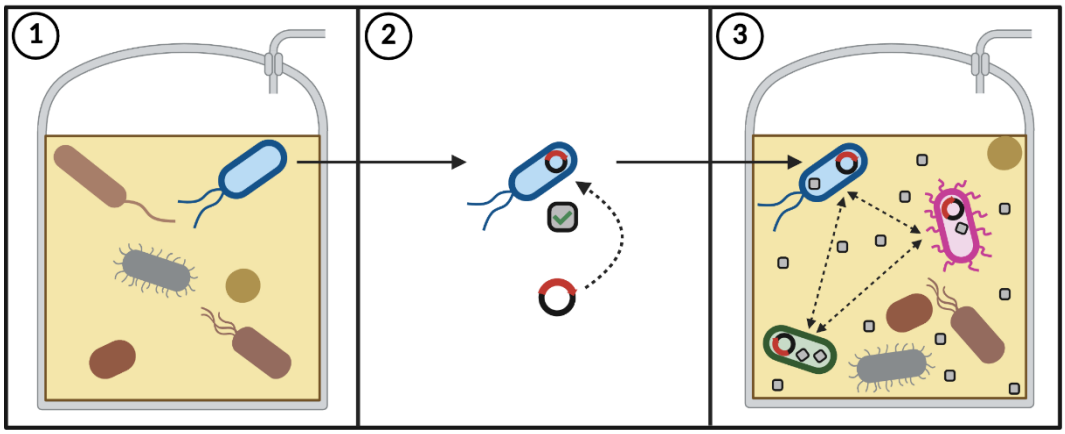


Figure 6.1: An example of combined approaches for AD community manipulation.
1) An organism with a desirable phenotype (blue) is isolated out from an AD community using selective plating approaches. **2)** Genetic tools ported from close laboratory relatives are used to impart the isolate with the ability to derive energy from a source (grey rectangle) not normally found in an anaerobic digester. **3)** The engineered organism (blue) and others reintroduced alongside it (pink, green) are even more stable in a mixed digester community supplemented with the new food source as endogenous organisms are unable to utilise it. Additional techniques such as engineered auxotrophy (dashes arrows) could further stabilise the desirable organisms as a mix.

6.4 Immediate further work

A key next step for the work reported in this project would be to isolate, store and identify more AD organisms, as individual organisms in pure culture represent the basis of most

of the techniques explored here. Culturomics approaches should be used to achieve this. Recent studies in gut systems such as the generation of a defined community of 104 bacterial species from culture collections (Cheng *et al.* 2022) and the integration of robotics pipelines to screen over 25,000 microbial isolates from human stool samples (Huang *et al.* 2023) demonstrate the great potential of this approach, providing a clear incentive for further culturomics studies focused on AD.

Improvements to the organism identification methods used in this project would represent a good candidate for further development as a high throughput and rapid way of screening isolates in an important component of successful culture based pipelines (Park *et al.* 2024). One possibility would be to introduce a MALDI-TOF step to rapidly screen newly isolated AD organisms (Howe *et al.* 2024), which would speed the process while simultaneously making it more economical, since full genome sequencing approaches like the pooling methods discussed in Chapter 4 could be reserved only for new organisms (i.e. not identified previously and therefore not already present in the library of organisms in storage).

6.5 Longer term further work

If bottom-up AD synthetic mixed communities like those described in this project could be produced, one possible future application could be as inocula for digester seeding and startup. Since different microbial consortia are likely to be better suited to different feedstocks and digester configurations, different mixes of seed inocula with defined microbiological compositions could be developed for different digester application (Xu *et al.* 2021; Dargode *et al.* 2022; Favale *et al.* 2025). If these mixes could be demonstrated to have superior functionality (for example enhanced stability and/or methane generation potential) it could be envisaged that they could be marketed for sale. The economic potential of this is demonstrated by AD additive products aimed at enhancing the function of existing digesters such as mixtures of trace elements and enzymes (Leca *et al.* 2023) which are already commercially available.

Another clear future application of an AD SynCom once developed would be as a scientific/research tool. AD communities are known to vary both within digesters over time and between different digesters (Lee *et al.* 2012; Calusinska *et al.* 2018), which

means comparing results between studies is not straightforward. A single defined AD microbiome which could be shared between different research groups could enable improved standardisation (Jourdain and Gu 2025). A model rhizosphere community consisting of 16 soil microorganisms has been developed based on a similar principle (Coker *et al.* 2022). The development of standards and protocols for mixed microbial community dissemination from public strain repositories will be needed to streamline SynCom based research (Northen *et al.* 2024). Another important consideration is preservation, since a functional mixed community would need to be stabilised for (potentially long term) storage and distribution (Coker *et al.* 2022). Freezing of mixed microbial samples as investigated in this thesis may not be ideal for this application due to the requirements of a potentially costly uninterrupted cold chain, but alternative techniques such as freeze drying which enable preserved sample storage at ambient temperatures may be more suitable (Rockinger *et al.* 2021; Bensch *et al.* 2022). Most research regarding preservation of microbes has been done with pure cultures (Kerckhof *et al.* 2014), so further research focused on preservation of mixed cultures would be beneficial for this application.

A key consideration for AD SynComs intended for industrial application is the issue of functional redundancy. Simplified defined mixes might perform adequately for more consistent AD feedstocks such as byproducts from commercial food processing facilities (Mioduska *et al.* 2023), but may be less applicable to feedstock types that are known to be highly variable such as sewage sludge (Chiappero *et al.* 2021; Van Landuyt *et al.* 2026). In natural systems, functional redundancy is thought to provide contingency against process fluctuations (Venkiteshwaran *et al.* 2015; De Vrieze and Verstraete 2016), so synthetic communities with a lower degree of redundancy might not be expected to exhibit the same level of robustness. A key factor here is that much of the source of this instability, and therefore need for redundancy, is the inherent variability/unpredictability of the system, much of which is thought to be caused by fluctuations in the feedstock (Theuerl *et al.* 2019; Ibarra-Esparza *et al.* 2025). Comprehensive monitoring of AD feed characteristics is generally limited at full scale plants (Wu *et al.* 2019), and if this could be improved, for example with advanced monitoring techniques, a more predictable and consistently composed system could potentially be generated, which means the underlying AD community may not need be as robust. A further possibility that improved

feedstock and process monitoring might enable is the use of pre-emptive bioaugmentation, i.e. the proactive addition of stored microbial strains known to be well suited to or capable of counteracting the predetermined destabilising conditions before they take effect.

If a more stable system could not be generated, another option would be to rationally engineer functional redundancy into the AD SynCom. A bottom-up generated microbial community of over 100 organisms has recently been demonstrated in a gut system (Cheng *et al.* 2022) so it would likely be feasible to generate an AD community with large numbers of contingency organisms. Such a community may be advantageous over non-defined communities as it would be understood at the organism level so could for example be further rationally refined. A key current limitation is that this kind of intervention would require an extremely detailed mechanistic knowledge of AD communities, so while unlikely feasible currently, it may be possible in the future, especially given the considerable advances that have been made in our understanding of this field in recent years (Lim *et al.* 2020; Harirchi *et al.* 2022; Zhang *et al.* 2024).

Directed evolution was explored in this project to improve the performance of one organism with respect to a single condition. A similar strategy employed in laboratory directed evolution is to incrementally increase the level of exposure to the factor of interest, which allows a progressive improvement in performance of the microorganism (Van den Bergh *et al.* 2018). For example, a population of *E. coli* with an 870-fold increased level of resistance to the antibiotic chloramphenicol was generated using this stepped approach (Toprak *et al.* 2011). Acclimation studies using a similar stepped exposure principle with the aim of enhancing the tolerance of AD communities to conditions of interest such as high levels of phenol (Madigou *et al.* 2016), ammonia (Bucci *et al.* 2023) and toxic pyrolysis byproducts (Zhou *et al.* 2019) have recently been reported. An advantage of traditional laboratory evolution approaches that involve the use of a single genetically defined organism as a starting point is that sequencing can be used to investigate the genetic basis of any evolved traits in the resulting higher performance derivative (Van den Bergh *et al.* 2018). A potential benefit of acclimation studies using mixed communities is there may be increased potential for performance improvements since different organisms are able to work synergistically with each other (Lawrence *et al.*

2012; Rodríguez Amor and Dal Bello 2019). These studies tend to start with undefined mixed microbial assemblages (Madigou *et al.* 2016; Zhou *et al.* 2019; Bucci *et al.* 2023), so one possible avenue for future research would be to use a defined SynCom of isogenic microbial strains as the starting point for acclimation, which would combine the strengths of both approaches by allowing a collection or genetically defined organisms to be evolved together.

The potential use of genetic modification to impart novel functionalities to microbes for use in AD was considered in this thesis. The problem of emerging contaminants found in AD such as antibiotics and microplastics has received increased attention recently (Harirchi *et al.* 2026). A potentially useful avenue for whole cell microbial biosensor development would be to generate strains able to report on or even degrade these substances inside digesters (Tang 2023). Genetically engineered microbial remediation systems intended for use in wastewater have been explored for plastic degradation (Yip *et al.* 2024) and antibiotic removal (Mu *et al.* 2018) in proof of concept studies. While microbial genetic modification holds obvious potential solutions to environmental challenges, including in AD, biocontainment (i.e. the prevention of release and propagation of genetically modified microorganisms or their DNA in unintended environments) is of crucial importance, both in terms of public acceptance and safety (Stirling and Silver 2020). Numerous approaches including genetic kill switches, engineered auxotrophy and physical encapsulation have been investigated as biocontainment strategies (Fuerte-Stone and Mimee 2026), but a key issue is that they inherently apply a strong selective pressure for organisms to circumvent or "escape" them (Stirling and Silver 2020). Combinatorial approaches using multiple containment techniques together can reduce the frequency of escapees (Gallagher *et al.* 2015), but the recent demonstration that escape rates can be considerably higher in real world conditions (i.e. outside of a controlled laboratory environment) (Hartig *et al.* 2024) demonstrates that there is still work to be done before complete biocontainment can be guaranteed (Stirling and Silver 2020).

Appendix:

Table A1: Species identified for investigation.
List of species belonging to the genera *Pseudomonas*, *Bacteroides*, *Bacillus* and *Clostridium* identified for investigation. Source indicates from where the species was obtained: MiDAS = the MiDAS database. Review = a review paper (see methods). Web = internet search.

Genus	Species	Source	Anaerobic?	Genetic Tools?
Pseudomonas	*mendocina*	MiDAS	No	Undetermined
Pseudomonas	*caeni*	MiDAS	Yes, facultative	No
Pseudomonas	*aeruginosa*	MiDAS	Yes, facultative	Yes
Pseudomonas	*alcaligenes*	Web	No	Undetermined
Pseudomonas	*pseudoalcaligenes*	Web	No	Undetermined
Pseudomonas	*resinovorans*	Web	Undetermined	Undetermined
Pseudomonas	*veronii*	Web	No	Undetermined
Pseudomonas	*putida*	Web	No	Undetermined
Pseudomonas	*stutzeri*	Web	Yes, facultative	Yes
Pseudomonas	*fluorescens*	Web	No	Undetermined
Pseudomonas	*aureofaciens*	Web	No	Undetermined
Pseudomonas	*protegenes*	Web	No	Undetermined
Pseudomonas	*chlororaphis*	Web	No	Undetermined
Pseudomonas	*aurantiaca*	Web	No	Undetermined
Pseudomonas	*oryzihabitans*	Web	No	Undetermined
Pseudomonas	*plecoglossicida*	Web	Yes	No
Pseudomonas	*syringae*	Web	No	Undetermined
Pseudomonas	*migulae*	Web	No	Undetermined
Pseudomonas	*graminis*	Web	No	Undetermined
Pseudomonas	*entemophila*	Review	No	Undetermined
Bacteroides	*graminisolvens*	MiDAS	Yes, obligate	No
Bacteroides	*dorei*	MiDAS	Yes, obligate	No
Bacteroides	*reticulotermitis*	MiDAS	Yes, obligate	No
Bacteroides	*uniformis*	MiDAS	Yes, obligate (aerotolerant)	Yes
Bacteroides	*massiliensis*	MiDAS	Yes, obligate	No
Bacteroides	*thetaiotaomicron*	MiDAS	Yes, obligate (aerotolerant)	Yes
Bacteroides	*caccae*	MiDAS	Yes, obligate	No
Bacteroides	*ovatus*	MiDAS	Yes, obligate (aerotolerant)	Yes
Bacteroides	*xylanisolvens*	MiDAS	Yes, obligate	Yes (limited)
Bacteroides	*vulgatus*	MiDAS	Yes, obligate (aerotolerant)	Yes
Bacteroides	*eggerthii*	MiDAS	Yes, obligate (aerotolerant)	Yes
Bacteroides	*coprocola*	MiDAS	Yes, obligate	No
Bacteroides	*plebeius*	MiDAS	Yes, obligate	No

Bacteroides	*stercoris*	MiDAS	Yes, obligate	No
Bacteroides	*cellulosilyticus*	MiDAS	Yes, obligate	Yes (limited)
Bacteroides	*luti*	MiDAS	Yes, obligate	No
Bacteroides	*faecis*	MiDAS	Yes, obligate	No
Bacteroides	*intestinalis*	MiDAS	Yes, obligate	No
Bacteroides	*paurosaccharolyticus*	MiDAS	Yes, obligate	No
Bacteroides	*finegoldii*	MiDAS	Yes, obligate	No
Bacteroides	*heparinolyticus*	MiDAS	Yes, obligate	No
Bacteroides	*coprophilus*	MiDAS	Yes, obligate	No
Bacteroides	*clarus*	MiDAS	Yes, obligate	No
Bacteroides	*salyersiae*	MiDAS	Yes, obligate	No
Bacteroides	*nordii*	MiDAS	Yes, obligate	No
Bacteroides	*acidifaciens*	Review	Yes, obligate	No
Bacteroides	*coprosuis*	Review	Yes, obligate	No
Bacteroides	*fragilis*	Review	Yes, obligate (aerotolerant)	Yes
Bacteroides	*helcogenes*	Review	Yes, obligate	No
Bacteroides	*salyersiae*	Review	Yes, obligate	No
Bacteroides	*pyogenes*	Review	Yes, obligate	No
Bacteroides	*goldsteinii*	Review	Yes, obligate	No
Bacteroides	*johnsonii*	Review	Yes, obligate	No
Bacillus	*psychrosaccharolyticus*	MiDAS	Yes, facultative	No
Bacillus	*thermophilus*	MiDAS	Undetermined	Undetermined
Bacillus	*subtilis*	Review	Yes, facultative	Yes
Bacillus	*licheniformis*	Review	Yes, facultative	Yes
Bacillus	*megaterium*	Review	Yes, facultative	Yes
Bacillus	*thuringiensis*	Review	Yes, facultative	Yes
Bacillus	*clausii*	Review	Yes	No
Bacillus	*pumilus*	Review	Yes, facultative	Yes
Bacillus	*amyloliquefaciens*	Review	Yes, facultative	Yes
Clostridium	*Spiroforme*	MiDAS	Yes, obligate	No
Clostridium	*disporicum*	MiDAS	Yes, obligate	No
Clostridium	*saudiense*	MiDAS	Yes, obligate	No
Clostridium	*perfringens*	MiDAS	Yes, obligate (aerotolerant)	Yes
Clostridium	*butyricum*	MiDAS	Yes, obligate	Yes (limited)
Clostridium	*paraputrificum*	MiDAS	Yes, obligate	Yes (limited)
Clostridium	*chromiireducens*	MiDAS	Yes, obligate	No
Clostridium	*saccharoperbutylacetonicum*	MiDAS	Yes, obligate	Yes
Clostridium	*beijerinckii*	MiDAS	Yes, obligate (some strains aerotolerant)	Yes
Clostridium	*aurantibutyricum*	MiDAS	Yes, obligate	No
Clostridium	*chartatabidum*	MiDAS	Yes, obligate	No
Clostridium	*bowmanii*	MiDAS	Yes, obligate	No

Clostridium	*senegalense*	MiDAS	Yes, obligate	No
Clostridium	*amazonense*	MiDAS	Yes, obligate	No
Clostridium	*thermopalmarium*	MiDAS	Yes, obligate	No
Clostridium	*novyi*	MiDAS	Yes, obligate	Yes
Clostridium	*cochlearium*	MiDAS	Yes, obligate	No
Clostridium	*frigidicarnis*	MiDAS	Yes, obligate	No
Clostridium	*fimetarium*	MiDAS	Yes, obligate	No
Clostridium	*algidixylanolyticum*	MiDAS	Yes, obligate	No
Clostridium	*thermoalcaliphium*	MiDAS	Yes, obligate	No
Clostridium (Clostridoides)	*difficile*	Review	Yes, obligate	Yes
Clostridium	*thermocellum*	Review	Yes, obligate	Yes (thermophile)
Clostridium	*tyrobutyricum*	Review	Yes, obligate	Yes
Clostridium	*acetobutylicum*	Review	Yes, obligate	Yes
Clostridium	*cellulolyticum*	Review	Yes, obligate	Yes
Clostridium	*ljunghdalii*	Review	Yes, obligate	Yes
Clostridium	*pasteurianium*	Review	Yes, obligate	Yes
Clostridium	*tetani*	Review	Yes, obligate	Yes
Clostridium	*botulinum*	Review	Yes, obligate	Yes
Clostridium	*sporogenes*	Review	Yes, obligate	Yes
Clostridium	*autoethanogenum*	Review	Yes, obligate	Yes
Clostridium	*cellulovorans*	Review	Yes, obligate	Yes (limited)
Clostridium	*saccharobutylicum*	Review	Yes, obligate	Yes (limited)
Clostridium	*Septicum*	Review	Yes, obligate	Yes (limited)
Clostridium	*Carboxidivorans*	Review	Yes, obligate	Yes (limited)
Clostridium	*Kluyveri*	Review	Yes, obligate	No
Clostridium	*Phytofermentans*	Review	Yes, obligate	Yes (pathogenic)
Clostridium (Paeniclostridium)	*Sordelli*	Review	Yes, obligate	Yes (pathogenic)

Table A2: List of species identified as "desirable" organisms.
The subset of the original list of 101 species (Table A1, Appendix) confirmed to be both anaerobic and have available genetic tools and subsequently selected. These form the basis for searches in metagenomic data for wild relatives.

Genus	Species	Genome Used	Genome Quality
Pseudomonas	*aeruginosa*	NC_002516.2	Reference
Pseudomonas	*stutzeri*	NZ_AP024722.1	Representative
Bacteroides	*uniformis*	NZ_CP072255.1	Representative
Bacteroides	*thetaiotaomicron*	NZ_CP040530.1	Representative
Bacteroides	*vulgatus*	NZ_CP072234.1	Representative
Bacteroides	*fragilis*	NZ_CP069563.1	Representative
Bacillus	*subtilis*	NC_000964.3	Reference
Bacillus	*licheniformis*	NZ_CP014842.1	Representative
Bacillus	*megaterium*	NZ_CP058255.1	Not Representative
Bacillus	*pumilus*	NZ_LT906438.1	Not Representative
Bacillus	*amyloliquefaciens*	NC_020272.1	Representative
Clostridium	*perfringens*	NC_008261.1	Representative
Clostridium	*saccharoperbutylacetonicum*	NC_020291.1	Representative
Clostridium	*beijerinckii*	NZ_CP073653.1	Representative
Clostridium (Clostridoides)	*difficile*	NZ_CP076401.1	Representative
Clostridium	*tyrobutyricum*	NZ_CP014170.1	Representative
Clostridium	*acetobutylicum*	NC_015687.1	Representative
Clostridium	*cellulolyticum*	NC_011898.1	Representative
Clostridium	*ljunghdalii*	NC_014328.1	Representative
Clostridium	*pasteurianium*	NZ_CP013019.1	Representative
Clostridium	*sporogenes*	NZ_CP011663.1	Representative
Clostridium	*autoethanogenum*	NZ_CP012395.1	Representative

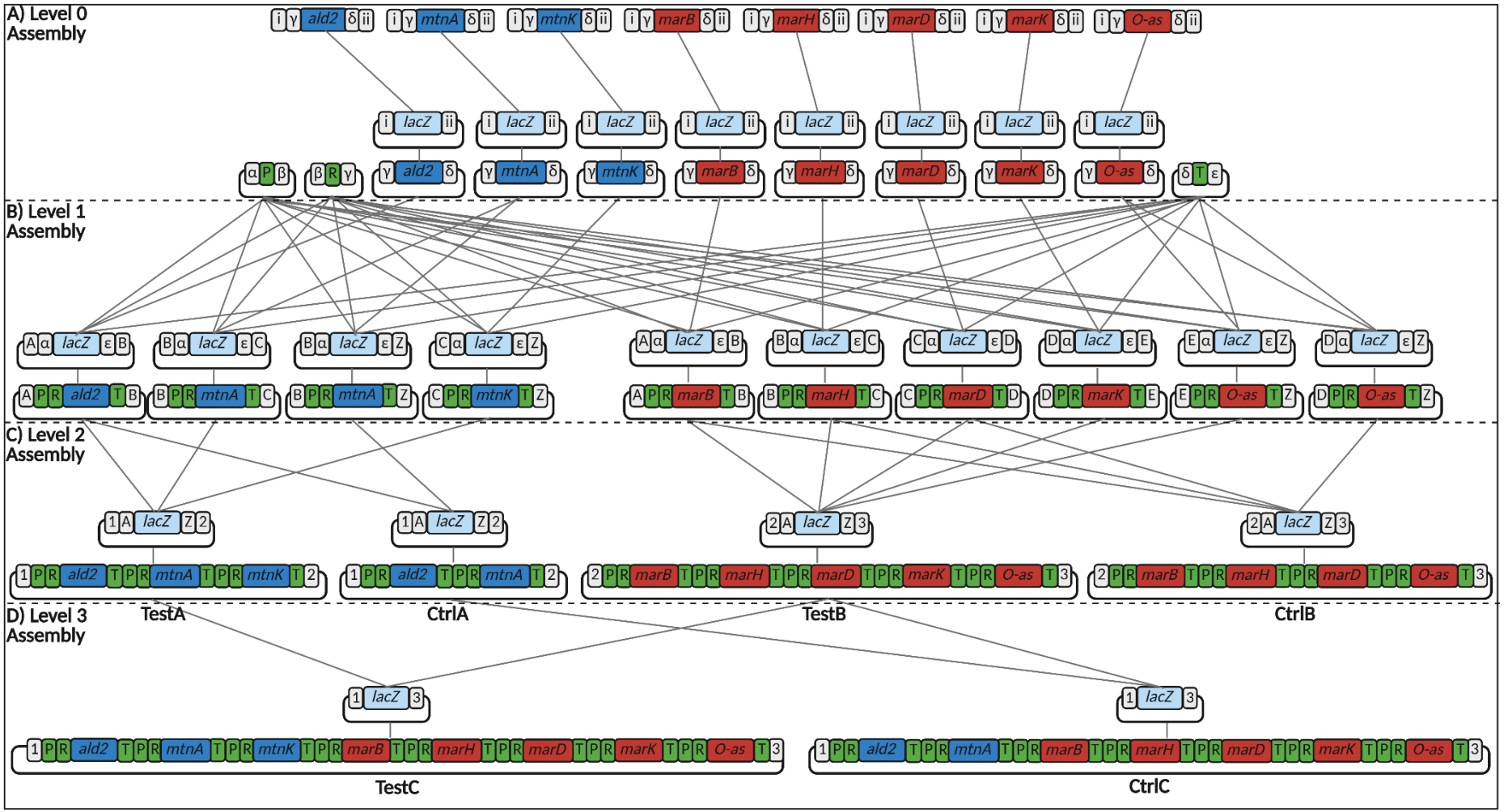


Figure A1: Ethylene and 2(methylthio)ethanol producing plasmid assembly scheme.

A series of Golden Gate reactions using the synthesised genes and Start-Stop kit components are used to sequentially assemble the three plasmids for producing ethylene and 2(methylthio)ethanol and their respective controls. **A)** Adapter sequences are added to the synthesised DNA fragments prior to ordering with inward facing SapI sites (γ, δ) nested within inward basing BsaI sites (i, ii). The level 0 holding vector pStA0 contains outward facing BsaI sites. A Golden Gate reaction with BsaI integrates the synthesised gene into the vector. Replacement of the *lacZ* gene enables blue/white screening of successful

ligations. **B)** Inward facing SapI sites are present on the level 0 vectors containing the eight coding sequences produced in panel A and on the level 0 vectors containing regulatory elements provided with the kit (α, β, γ, δ, ε). The level 1 holding vectors (e.g. pStA1AB) contain outward facing SapI sites. A Golden Gate reaction with SapI integrates the coding sequences and regulatory elements in the correct order into the desired level 1 vector to produce a transcriptional unit. Only the integration of *ald2* into pStA1AB is shown, but this step is conducted for all eight genes with appropriate level 1 vectors. **C)** Inward facing BsaI sites are present on each of the level 1 vectors containing the transcriptional units produced in panel B (A, B, C, D, E, Z). The level 2 holding vectors (pStA212 and pStA213) contain outward facing BsaI sites. A Golden Gate reaction with BsaI integrates the transcriptional units *ald2*, *mtnA* and *mtnK* into pStA212 producing plasmid TestA, which is designed to produce 2(methylthio)ethanol in commensal *E. coli*. A separate Golden Gate reaction with BsaI integrates the transcriptional units *marBHDK* and *O-as* into pStA213 producing plasmid TestB, which is designed to produce ethylene in ExPEC *E. coli*. Regulatory elements are not shown. Controls plasmids CtrlA and CtrlB are produced using a similar system. **D)** Inward facing BbsI sites are present on the level 2 vectors containing the transcriptional units *ald2-mtnA-mtnK* and *marBDHK-O-as*. Outward facing BbsI sites on the level 3 holding vector pStA313 allows assembly of all units into a single plasmid, which is designed to produce ethylene in commensal *E. coli*. Regulatory elements are not shown. Green boxes: regulatory elements. P = promoter. RBS = ribosome binding site. T = terminator. Blue and red boxes: genes. Numbered and lettered grey boxes: Golden Gate overlap sites. Grey lines indicate components combined to produce the next level of plasmids in the scheme.

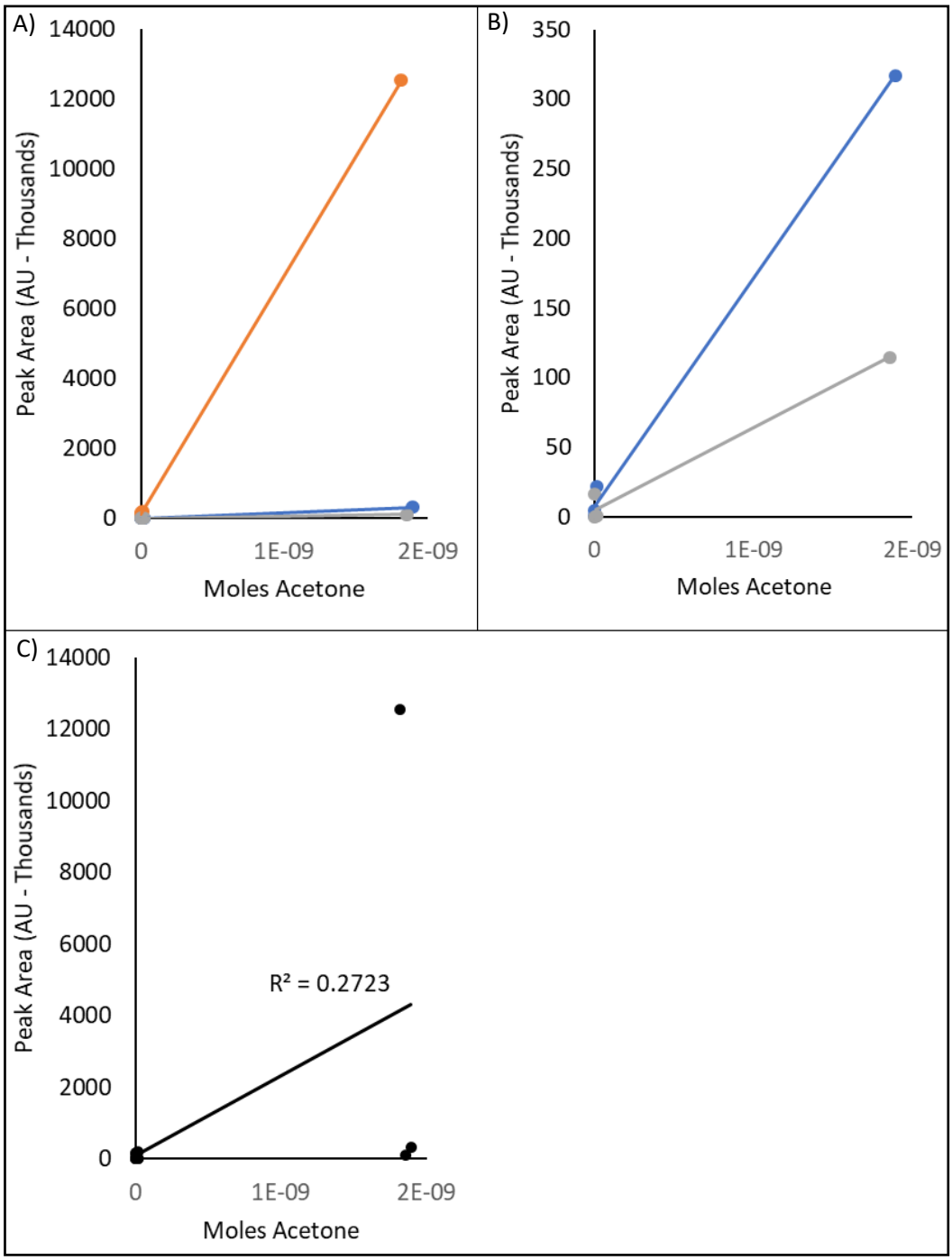


Figure A2: Acetone calibrations.
A) Four-point acetone calibration curves taken on three separate days plotted separately. **B)** The two curves with the lower peak areas plotted separately for clarity. **C)** All data points plotted together. The coefficient of determination (R^2) is indicated above the calibration curve. Note that because the minimum dilution factor which could be achieved for standards on this GC-MS was 100-fold, the lower two acetone concentrations are included but obscured by the two higher concentrations in these graphs. AU = arbitrary units.

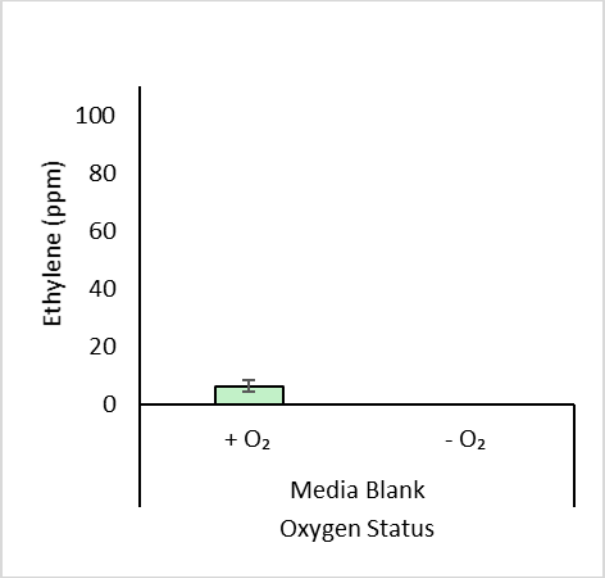


Figure A3: Initial ethylene test media controls.
Headspace ethylene measurements for media only controls. Point are means of triplicate vessels. Error bars are +/- one standard deviation.

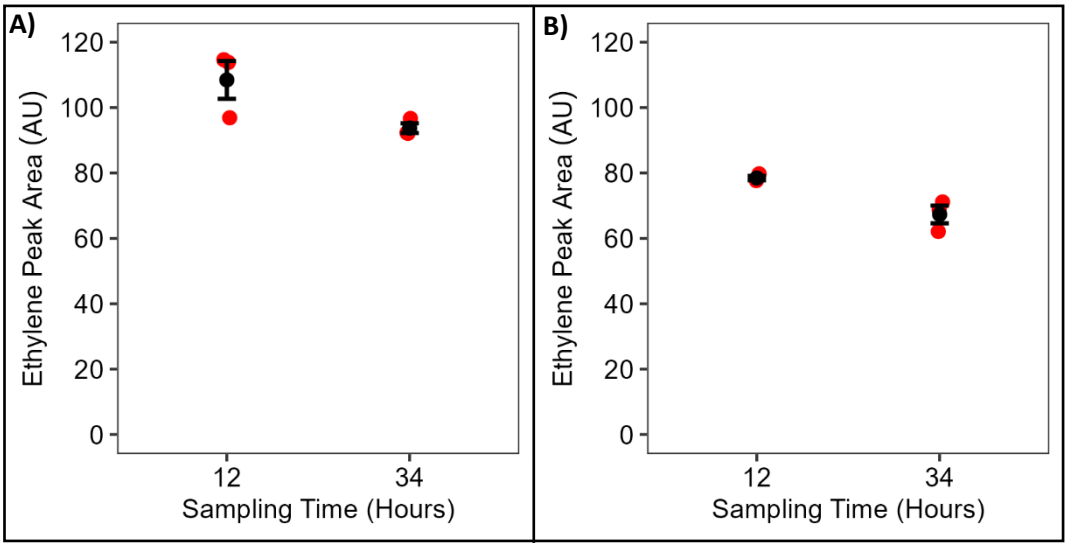


Figure A4: Ethylene headspace repeat readings.
Triplicate headspace ethylene readings for cultures TestC A3 and CtrlC B6 at two different sampling timepoints. **A)** TestC A3 (aerobic conditions). **B)** CtrlC B6 (anaerobic conditions). Points are means of triplicate 1 mL headspace samples taken from the same vessel. Error bars are +/- one standard error.

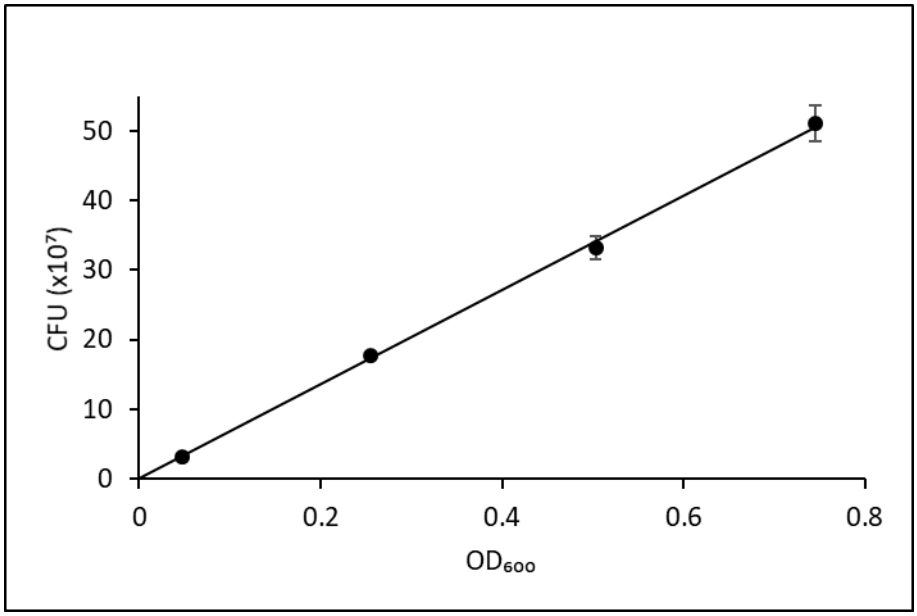


Figure A5: CFU to OD_{600} calibrations for *E. coli* RFP strain.
Points are mean colony forming units (CFU) calculated from triplicate platings at four different starting OD_{600} readings. Error bars are +/- one standard deviation. Very small error bars for two of the points are not visible. Note counts for this experiment exceeded 300 colonies per plate.

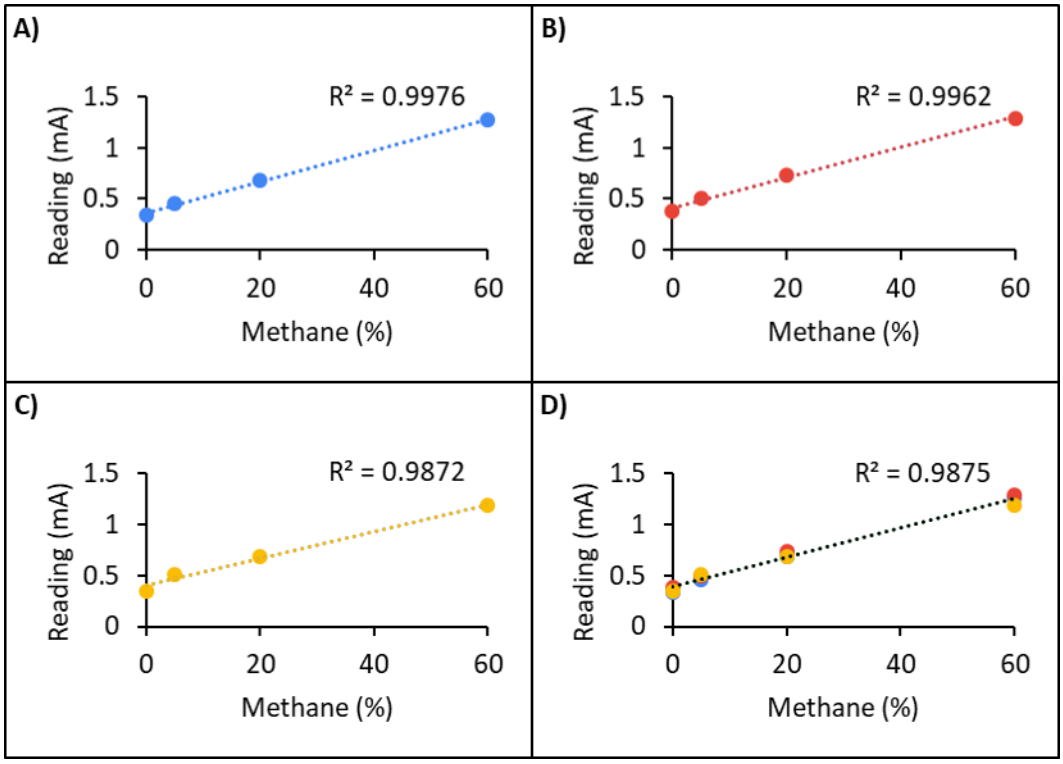


Figure A6: Reproducible methane content measurements with an IR sensor.
Four-point methane calibration curves generated using the IR sensor. A, B, C) Calibrations generated on three separate days. D) The three days plotted together. The coefficient of determination (R^2) is indicated above the calibration curve in each case.

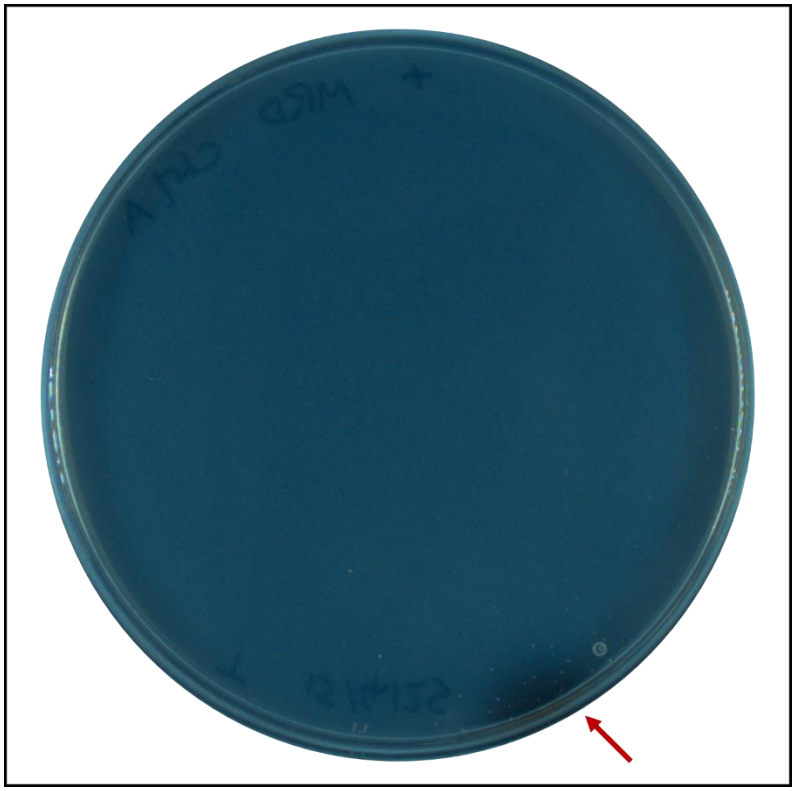


Figure A7: Irregularly shaped colonies on T and S plates.
Representative image of the irregularly shaped colonies thought to be contamination recovered on T and S plates. These colonies are most visible in the area of discoloured agar (red arrow).

Reference List

Achtman, M. and Wagner, M. (2008) 'Microbial diversity and the genetic nature of microbial species', *Nature Reviews Microbiology*, 6(6), 431-40, available: http://dx.doi.org/10.1038/nrmicro1872.

Adams, B.L. (2016) 'The Next Generation of Synthetic Biology Chassis: Moving Synthetic Biology from the Laboratory to the Field', *ACS Synthetic Biology*, 5(12), 1328-1330, available: http://dx.doi.org/10.1021/acssynbio.6b00256.

Aldridge, J., Carr, S., Weber, K.A. and Buan, N.R. (2021) 'Anaerobic Production of Isoprene by Engineered Methanosarcina Species Archaea', *Applied and Environmental Microbiology*, 87(6), available: http://dx.doi.org/10.1128/aem.02417-20.

Alengebawy, A., Ran, Y., Osman, A.I., Jin, K., Samer, M. and Ai, P. (2024) 'Anaerobic digestion of agricultural waste for biogas production and sustainable bioenergy recovery: a review', *Environmental Chemistry Letters*, 22(6), 2641-2668, available: http://dx.doi.org/10.1007/s10311-024-01789-1.

Alessi, A.M., Redeker, K.R. and Chong, J.P.J. (2018) 'A practical introduction to microbial molecular ecology through the use of isolation chips', *Ecology and Evolution*, 8(24), 12286-12298, available: http://dx.doi.org/10.1002/ece3.4748.

Ali, S.A., Mittal, D. and Kaur, G. (2021) 'In-situ monitoring of xenobiotics using genetically engineered whole-cell-based microbial biosensors: recent advances and outlook', *World Journal of Microbiology and Biotechnology*, 37(5), 81, available: http://dx.doi.org/10.1007/s11274-021-03024-3.

Alipoursarbani, M., Tideman, J., López, M. and Abendroth, C. (2026) 'Bioaugmentation in anaerobic digesters: a systematic review', *Biotechnology for Biofuels and Bioproducts*, 19(1), 25, available: http://dx.doi.org/10.1186/s13068-026-02746-6.

Allali, I., Arnold, J.W., Roach, J., Cadenas, M.B., Butz, N., Hassan, H.M., Koci, M., Ballou, A., Mendoza, M., Ali, R. and Azcarate-Peril, M.A. (2017) 'A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome', *BMC Microbiology*, 17(1), 194, available: http://dx.doi.org/10.1186/s12866-017-1101-8.

Allen, A. (2001) 'Containment landfills: the myth of sustainability', *Engineering Geology*, 60(1), 3-19, available: http://dx.doi.org/10.1016/S0013-7952(00)00084-3.

Alonso-Gutierrez, J., Chan, R., Batth, T.S., Adams, P.D., Keasling, J.D., Petzold, C.J. and Lee, T.S. (2013) 'Metabolic engineering of Escherichia coli for limonene and perillyl alcohol production', *Metabolic Engineering*, 19, 33-41, available: http://dx.doi.org/10.1016/j.ymben.2013.05.004.

Alshammari, M., Ahmad, A., AlKhulaifi, M., Al Farraj, D., Alsudir, S., Alarawi, M., Takashi, G. and Alyamani, E. (2023) 'Reduction of biofilm formation of Escherichia coli by targeting quorum sensing and adhesion genes using the CRISPR/Cas9-HDR approach, and its clinical application on urinary catheter', *Journal of Infection and Public Health*, 16(8), 1174-1183, available: http://dx.doi.org/10.1016/j.jiph.2023.05.026.

Amekan, Y. (2021) *Identifying Metabolically Active Microbial Populations in Anaerobic Digestion Through Bioorthogonal Non-Canonical Amino Acid Tagging and Affinity-Based Cell Separation*, unpublished thesis (PhD), University of York.

Amha, Y.M., Anwar, M.Z., Brower, A., Jacobsen, C.S., Stadler, L.B., Webster, T.M. and Smith, A.L. (2018) 'Inhibition of anaerobic digestion processes: Applications of molecular tools', *Bioresource Technology*, 247, 999-1014, available: http://dx.doi.org/10.1016/j.biortech.2017.08.210.

André, A.C., Debande, L. and Marteyn, B.S. (2021) 'The selective advantage of facultative anaerobes relies on their unique ability to cope with changing oxygen levels during infection', *Cellular Microbiology*, 23(8), e13338, available: http://dx.doi.org/10.1111/cmi.13338.

Appels, L., Baeyens, J., Degrève, J. and Dewil, R. (2008) 'Principles and Potential of the Anaerobic Digestion of Waste-Activated Sludge', *Progress in Energy and Combustion Science*, 34, 755-781, available: http://dx.doi.org/10.1016/j.pecs.2008.06.002.

Arbefeville, S.S., Timbrook, T.T. and Garner, C.D. (2024) 'Evolving strategies in microbe identification-a comprehensive review of biochemical, MALDI-TOF MS and molecular testing methods', *Journal of Antimicrobial Chemotherapy*, 79(12 Suppl 2), i2-i8, available: http://dx.doi.org/10.1093/jac/dkae275.

Ariunbaatar, J., Panico, A., Esposito, G., Pirozzi, F. and Lens, P.N.L. (2014) 'Pretreatment methods to enhance anaerobic digestion of organic solid waste', *Applied Energy*, 123, 143-156, available: http://dx.doi.org/10.1016/j.apenergy.2014.02.035.

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

Astals, S., Koch, K., Weinrich, S., Hafner, S.D., Tait, S. and Peces, M. (2020) 'Impact of Storage Conditions on the Methanogenic Activity of Anaerobic Digestion Inocula', *Water*, 12(5), available: http://dx.doi.org/10.3390/w12051321.

Atelge, M.R., Krisa, D., Kumar, G., Eskicioglu, C., Nguyen, D.D., Chang, S.W., Atabani, A.E., Al-Muhtaseb, A.H. and Unalan, S. (2020) 'Biogas Production from Organic Waste: Recent Progress and Perspectives', *Waste and Biomass Valorization*, 11(3), 1019-1040, available: http://dx.doi.org/10.1007/s12649-018-00546-0.

Avrani, S., Katz, S. and Hershberg, R. (2020) 'Adaptations Accumulated under Prolonged Resource Exhaustion Are Highly Transient', *mSphere*, 5(4), e00388-20, available: http://dx.doi.org/10.1128/mSphere.00388-20.

Babu, M. and Snyder, M. (2023) 'Multi-Omics Profiling for Health', *Molecular & Cellular Proteomics*, 22(6), 100561, available: http://dx.doi.org/10.1016/j.mcpro.2023.100561.

Bai, G., Yin, Y., Zhang, Z., Li, Y., Liu, Y., Wang, R., He, Z.-Q., Han, C., Zhang, R., Cui, J.-Q., Li, B.-Z. and Bai, H. (2025) 'Bioaugmentation strategies for methane production from digestate-alkali pretreated biomass via anaerobic digestion', *Bioresources and Bioprocessing*, 12(1), 148, available: http://dx.doi.org/10.1186/s40643-025-00990-6.

Barber, W.P.F. (2016) 'Thermal hydrolysis for sewage treatment: A critical review', *Water Research*, 104, 53-71, available: http://dx.doi.org/10.1016/j.watres.2016.07.069.

Barker Clive, S., Prüß Birgit, M. and Matsumura, P. (2004) 'Increased Motility of Escherichia coli by Insertion Sequence Element Integration into the Regulatory Region of the flhD Operon', *Journal of Bacteriology*, 186(22), 7529-7537, available: http://dx.doi.org/10.1128/jb.186.22.7529-7537.2004.

Basak, B., Patil, S.M., Saha, S., Kurade, M.B., Ha, G.-S., Govindwar, S.P., Lee, S.S., Chang, S.W., Chung, W.J. and Jeon, B.-H. (2021) 'Rapid recovery of methane yield in organic overloaded-failed anaerobic digesters through bioaugmentation with acclimatized microbial consortium', *Science of The Total Environment*, 764, 144219, available: http://dx.doi.org/10.1016/j.scitotenv.2020.144219.

Behringer, M.G., Choi, B.I., Miller, S.F., Doak, T.G., Karty, J.A., Guo, W. and Lynch, M. (2018) 'Escherichia coli cultures maintain stable subpopulation structure during long-term evolution', *Proceedings of the National Academy of Sciences of the United States of America*, 115(20), E4642-E4650, available: http://dx.doi.org/10.1073/pnas.1708371115.

Bellali, S., Lo, C.I., Naud, S., Fonkou, M.D.M., Armstrong, N., Raoult, D., Fournier, P.E. and Fenollar, F. (2019) 'Parabacteroides massiliensis sp. nov., a new bacterium isolated from a fresh human stool specimen', *New Microbes and New Infections*, 32, 100602, available: http://dx.doi.org/10.1016/j.nmni.2019.100602.

Bennett, R.K., Gonzalez, J.E., Whitaker, W.B., Antoniewicz, M.R. and Papoutsakis, E.T. (2018) 'Expression of heterologous non-oxidative pentose phosphate pathway from Bacillus methanolicus and phosphoglucose isomerase deletion improves methanol assimilation and metabolite production by a synthetic Escherichia coli methylotroph', *Metabolic Engineering*, 45, 75-85, available: http://dx.doi.org/10.1016/j.ymben.2017.11.016.

Bensch, H.M., Tolf, C., Waldenström, J., Lundin, D. and Zöttl, M. (2022) 'Freeze-drying can replace cold-chains for transport and storage of fecal microbiome samples', *PeerJ*, 10, e13095, available: http://dx.doi.org/10.7717/peerj.13095.

Benton, P.M., Christiansen, J., Dean, D.R. and Seefeldt, L.C. (2001) 'Stereospecificity of acetylene reduction catalyzed by nitrogenase', *Journal of the American Chemical Society*, 123(9), 1822-7, available: http://dx.doi.org/10.1021/ja003662x.

Berde, M.S. and Jolis, D. (2018) 'Solubilization of Organics Due to Thermal Hydrolysis Pretreatment and the Shift in Microbial Population in Anaerobic Digestion', *Water Environment Research*, 90(8), 761-767, available: http://dx.doi.org/10.2175/106143017x15131012153194.

Bermejo, L.L., Welker, N.E. and Papoutsakis, E.T. (1998) 'Expression of Clostridium acetobutylicum ATCC 824 genes in Escherichia coli for acetone production and acetate detoxification', *Applied and Environmental Microbiology*, 64(3), 1079-85, available: http://dx.doi.org/10.1128/aem.64.3.1079-1085.1998.

Berner, D. and Viernstein, H. (2006) 'Effect of protective agents on the viability of Lactococcus lactis subjected to freeze-thawing and freeze-drying', *Scientia Pharmaceutica*, 74(3), 137-149, available: http://dx.doi.org/10.3797/scipharm.2006.74.137.

Bhadra, S., Paik, I., Torres, J.A., Fadanka, S., Gandini, C., Akligoh, H., Molloy, J. and Ellington, A.D. (2022) 'Preparation and Use of Cellular Reagents: A Low-resource Molecular Biology Reagent Platform', *Current Protocols*, 2(3), e387, available: http://dx.doi.org/10.1002/cpz1.387.

Bhattad, U., Venkiteshwaran, K., Cherukuri, K., Maki, J.S. and Zitomer, D.H. (2017) 'Activity of methanogenic biomass after heat and freeze drying in air', *Environmental Science: Water Research & Technology*, 3(3), 462-471, available: http://dx.doi.org/10.1039/C7EW00049A.

Blasco, L., Kahala, M., Ervasti, S. and Tampio, E. (2022) 'Dynamics of microbial community in response to co-feedstock composition in anaerobic digestion', *Bioresource Technology*, 364, 128039, available: http://dx.doi.org/10.1016/j.biortech.2022.128039.

Bolnick, D., Barrett, R., Oke, K., Rennison, D. and Stuart, Y. (2018) '(Non)Parallel Evolution', *Annual Review of Ecology, Evolution, and Systematics*, 49, available: http://dx.doi.org/10.1146/annurev-ecolsys-110617-062240.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J.E., Bittinger, K., Brejnrod, A., Brislawn, C.J., Brown, C.T., Callahan, B.J., Caraballo-Rodríguez, A.M., Chase, J., Cope, E.K., Da Silva, R., Diener, C., Dorrestein, P.C., Douglas, G.M., Durall, D.M., Duvallet, C., Edwardson, C.F., Ernst, M., Estaki, M., Fouquier, J., Gauglitz, J.M., Gibbons, S.M., Gibson, D.L., Gonzalez, A., Gorlick, K., Guo, J., Hillmann, B., Holmes, S., Holste, H., Huttenhower, C., Huttley, G.A., Janssen, S., Jarmusch, A.K., Jiang, L., Kaehler, B.D., Kang, K.B., Keefe, C.R., Keim, P., Kelley, S.T., Knights, D., Koester, I., Kosciolek, T., Kreps, J., Langille, M.G.I., Lee, J., Ley, R., Liu, Y.-X., Loftfield, E., Lozupone, C., Maher, M., Marotz, C., Martin, B.D., McDonald, D., McIver, L.J., Melnik, A.V., Metcalf, J.L., Morgan, S.C., Morton, J.T., Naimey, A.T., Navas-Molina, J.A., Nothias, L.F., Orchanian, S.B., Pearson, T., Peoples, S.L., Petras, D., Preuss, M.L., Pruesse, E., Rasmussen, L.B., Rivers, A., Robeson, M.S., Rosenthal, P., Segata, N., Shaffer, M., Shiffer, A., Sinha, R., Song, S.J., Spear, J.R., Swafford, A.D., Thompson, L.R., Torres, P.J., Trinh, P., Tripathi, A., Turnbaugh, P.J., Ul-Hasan, S., van der Hooft, J.J.J., Vargas, F., Vázquez-Baeza, Y., Vogtmann, E., von Hippel, M., Walters, W.*, et al.* (2019) 'Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2', *Nature Biotechnology*, 37(8), 852-857, available: http://dx.doi.org/10.1038/s41587-019-0209-9.

Bonk, F., Popp, D., Weinrich, S., Sträuber, H., Kleinsteuber, S., Harms, H. and Centler, F. (2018) 'Intermittent fasting for microbes: how discontinuous feeding increases functional stability in anaerobic digestion', *Biotechnology for Biofuels*, 11(1), 274, available: http://dx.doi.org/10.1186/s13068-018-1279-5.

Bonnet, M., Lagier, J.C., Raoult, D. and Khelaifia, S. (2020) 'Bacterial culture through selective and non-selective conditions: the evolution of culture media in clinical microbiology', *New Microbes and New Infections*, 34, 100622, available: http://dx.doi.org/10.1016/j.nmni.2019.100622.

Boynton, P.J., Stelkens, R., Kowallik, V. and Greig, D. (2017) 'Measuring microbial fitness in a field reciprocal transplant experiment', *Molecular Ecology Resources*, 17(3), 370-380, available: http://dx.doi.org/10.1111/1755-0998.12562.

Braga, P.A.C., Tata, A., Gonçalves dos Santos, V., Barreiro, J.R., Schwab, N.V., Veiga dos Santos, M., Eberlin, M.N. and Ferreira, C.R. (2013) 'Bacterial identification: from the agar plate to the mass spectrometer', *RSC Advances*, 3(4), 994-1008, available: http://dx.doi.org/10.1039/C2RA22063F.

Breitwieser, F.P. and Salzberg, S.L. (2020) 'Pavian: interactive analysis of metagenomics data for microbiome studies and pathogen identification', *Bioinformatics*, 36(4), 1303-1304, available: http://dx.doi.org/10.1093/bioinformatics/btz715.

Buan, N.R. (2018) 'Methanogens: pushing the boundaries of biology', *Emerging Topics in Life Sciences*, 2(4), 629-646, available: http://dx.doi.org/10.1042/etls20180031.

Bucci, L., Ghiotto, G., Zampieri, G., Raga, R., Favaro, L., Treu, L. and Campanaro, S. (2023) 'Adaptation of Anaerobic Digestion Microbial Communities to High Ammonium Levels: Insights from Strain-Resolved Metagenomics', *Environmental Science & Technology*, 58(1), 580-590, available: http://dx.doi.org/10.1021/acs.est.3c07737.

Bukin, Y.S., Galachyants, Y.P., Morozov, I.V., Bukin, S.V., Zakharenko, A.S. and Zemskaya, T.I. (2019) 'The effect of 16S rRNA region choice on bacterial community metabarcoding results', *Scientific Data*, 6, 190007, available: http://dx.doi.org/10.1038/sdata.2019.7.

Bytnerowicz, T.A., Min, E., Griffin, K.L. and Menge, D.N.L. (2019) 'Repeatable, continuous and real-time estimates of coupled nitrogenase activity and carbon exchange at the whole-plant scale', *Methods in Ecology and Evolution*, 10(7), 960-970, available: http://dx.doi.org/10.1111/2041-210X.13186.

Calero, P. and Nikel, P.I. (2019) 'Chasing bacterial chassis for metabolic engineering: a perspective review from classical to non-traditional microorganisms', *Microbial Biotechnology*, 12(1), 98-124, available: http://dx.doi.org/10.1111/1751-7915.13292.

Callahan, B.J., McMurdie, P.J. and Holmes, S.P. (2017) 'Exact sequence variants should replace operational taxonomic units in marker-gene data analysis', *The ISME Journal*, 11(12), 2639-2643, available: http://dx.doi.org/10.1038/ismej.2017.119.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J. and Holmes, S.P. (2016) 'DADA2: High-resolution sample inference from Illumina amplicon data', *Nature Methods*, 13(7), 581-3, available: http://dx.doi.org/10.1038/nmeth.3869.

Calloni, G., Chen, T., Schermann, S.M., Chang, H.C., Genevaux, P., Agostini, F., Tartaglia, G.G., Hayer-Hartl, M. and Hartl, F.U. (2012) 'DnaK functions as a central hub in the E. coli chaperone network', *Cell Rep*, 1(3), 251-64, available: http://dx.doi.org/10.1016/j.celrep.2011.12.007.

Calusinska, M., Goux, X., Fossépré, M., Muller, E.E.L., Wilmes, P. and Delfosse, P. (2018) 'A year of monitoring 20 mesophilic full-scale bioreactors reveals the existence of stable but different core microbiomes in bio-waste and wastewater anaerobic digestion systems', *Biotechnology for Biofuels*, 11(1), 196, available: http://dx.doi.org/10.1186/s13068-018-1195-8.

Calusinska, M., Marynowska, M., Goux, X., Lentzen, E. and Delfosse, P. (2016) 'Analysis of dsDNA and RNA viromes in methanogenic digesters reveals novel viral genetic diversity', *Environmental Microbiology*, 18(4), 1162-75, available: http://dx.doi.org/10.1111/1462-2920.13127.

Campanaro, S., Treu, L., Rodriguez, R.L., Kovalovszki, A., Ziels, R.M., Maus, I., Zhu, X., Kougias, P.G., Basile, A., Luo, G., Schlüter, A., Konstantinidis, K.T. and Angelidaki, I. (2020) 'New insights from the biogas microbiome by comprehensive genome-resolved metagenomics of nearly 1600 species originating from multiple anaerobic digesters', *Biotechnology for Biofuels*, 13, 25, available: http://dx.doi.org/10.1186/s13068-020-01679-y.

Campbell, R.P., Whittington, A.C., Zorio, D.A.R. and Miller, B.G. (2023) 'Recruitment of a Middling Promiscuous Enzyme Drives Adaptive Metabolic Evolution in Escherichia coli', *Molecular Biology and Evolution*, 40(9), msad202, available: http://dx.doi.org/10.1093/molbev/msad202.

Cansdale, A. (2020) *CLUSTard: An automated pipeline for metagenomic clustering using read abundance over time*, unpublished thesis (MSc), University of York.

Cansdale, A., Forrester, S., Innard, N., Amekan, Y., Head, C.R., Pulford, A., Saye, L., Reilly, M. and Chong, J.P.J. (2023) 'The Role of Biomethane in Reaching Net Carbon Zero' in Oncel, S. S., ed., *A Sustainable Green Future: Perspectives on Energy, Economy, Industry, Cities and Environment*, Cham: Springer International Publishing, 575-594.

Canziani, R. and Spinosa, L. (2019) 'Sludge from wastewater treatment plants' in Prasad, M. N. V., de Campos Favas, P. J., Vithanage, M. and Mohan, S. V., eds., *Industrial and Municipal Sludge*, Oxford: Butterworth-Heinemann, 3-30.

Cao, J., Zhang, C., Li, X., Wang, X., Dai, X. and Xu, Y. (2025) 'Microbial Community and Metabolic Pathways in Anaerobic Digestion of Organic Solid Wastes: Progress, Challenges and Prospects', *Fermentation*, 11(8), 457, available: http://dx.doi.org/10.3390/fermentation11080457.

Cao, X., Hamilton, J.J. and Venturelli, O.S. (2019) 'Understanding and Engineering Distributed Biochemical Pathways in Microbial Communities', *Biochemistry*, 58(2), 94-107, available: http://dx.doi.org/10.1021/acs.biochem.8b01006.

Capitani, G., Eidam, O., Glockshuber, R. and Grütter, M.G. (2006) 'Structural and functional insights into the assembly of type 1 pili from Escherichia coli', *Microbes and Infection*, 8(8), 2284-2290, available: http://dx.doi.org/10.1016/j.micinf.2006.03.013.

Carballa, M., Regueiro, L. and Lema, J.M. (2015) 'Microbial management of anaerobic digestion: exploiting the microbiome-functionality nexus', *Current Opinion in Biotechnology*, 33, 103-111, available: http://dx.doi.org/10.1016/j.copbio.2015.01.008.

Carr, S. and Buan, N.R. (2022) 'Insights into the biotechnology potential of Methanosarcina', *Frontiers in Microbiology*, 13, 1034671, available: http://dx.doi.org/10.3389/fmicb.2022.1034674.

Carrère, H., Dumas, C., Battimelli, A., Batstone, D.J., Delgenès, J.P., Steyer, J.P. and Ferrer, I. (2010) 'Pretreatment methods to improve sludge anaerobic degradability: A review', *Journal of Hazardous Materials*, 183(1), 1-15, available: http://dx.doi.org/10.1016/j.jhazmat.2010.06.129.

Castro, H., Queirolo, M., Quevedo, M. and Muxí, L. (2002) 'Preservation methods for the storage of anaerobic sludges', *Biotechnology Letters*, 24(4), 329-333, available: http://dx.doi.org/10.1023/A:1014080526608.

Cecchini, G., Schröder, I., Gunsalus, R.P. and Maklashina, E. (2002) 'Succinate dehydrogenase and fumarate reductase from Escherichia coli', *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1553(1), 140-157, available: http://dx.doi.org/10.1016/S0005-2728(01)00238-9.

Chang, A.Y., Chau, V., Landas, J.A. and Pang, Y. (2017) 'Preparation of calcium competent Escherichia coli and heat-shock transformation', *JEMI methods*, 1, 22-25.

Chang, S. and Cohen, S.N. (1979) 'High frequency transformation of Bacillus subtilis protoplasts by plasmid DNA', *Molecular and General Genetics*, 168(1), 111-115, available: http://dx.doi.org/10.1007/BF00267940.

Charubin, K., Bennett, R.K., Fast, A.G. and Papoutsakis, E.T. (2018) 'Engineering Clostridium organisms as microbial cell-factories: challenges & opportunities', *Metabolic Engineering*, 50, 173-191, available: http://dx.doi.org/10.1016/j.ymben.2018.07.012.

Chen, J., Wade, M.J., Dolfing, J. and Soyer, O.S. (2019) 'Increasing sulfate levels show a differential impact on synthetic communities comprising different methanogens and a sulfate reducer', *Journal of the Royal Society Interface*, 16(154), 20190129, available: http://dx.doi.org/10.1098/rsif.2019.0129.

Chen, L., Li, S., Zhang, J., Zhen, F., Shang, Z., Yan, M., Zhang, Y., Zhang, P., Sun, Y. and Li, Y. (2025) 'A critical review on bioaugmentation assisted anaerobic digestion for methane production: Performances, microbiome-functionalities and challenges', *Journal of Environmental Management*, 380, 125127, available: http://dx.doi.org/10.1016/j.jenvman.2025.125127.

Chen, Y., Cheng, J.J. and Creamer, K.S. (2008) 'Inhibition of anaerobic digestion process: A review', *Bioresource Technology*, 99(10), 4044-4064, available: http://dx.doi.org/10.1016/j.biortech.2007.01.057.

Cheng, A.G., Ho, P.-Y., Aranda-Díaz, A., Jain, S., Yu, F.B., Meng, X., Wang, M., Iakiviak, M., Nagashima, K., Zhao, A., Murugkar, P., Patil, A., Atabakhsh, K., Weakley, A., Yan, J., Brumbaugh, A.R., Higginbottom, S., Dimas, A., Shiver, A.L., Deutschbauer, A., Neff, N., Sonnenburg, J.L., Huang, K.C. and Fischbach, M.A. (2022) 'Design, construction, and in vivo augmentation of a complex gut microbiome', *Cell*, 185(19), 3617-3636.e19, available: http://dx.doi.org/10.1016/j.cell.2022.08.003.

Cheng, H.-Y., Masiello, C.A., Bennett, G.N. and Silberg, J.J. (2016) 'Volatile Gas Production by Methyl Halide Transferase: An In Situ Reporter Of Microbial Gene

Expression In Soil', *Environmental Science & Technology*, 50(16), 8750-8759, available: http://dx.doi.org/10.1021/acs.est.6b01415.

Cheng, H.-Y., Masiello, C.A., Del Valle, I., Gao, X., Bennett, G.N. and Silberg, J.J. (2018) 'Ratiometric Gas Reporting: A Nondisruptive Approach To Monitor Gene Expression in Soils', *ACS Synthetic Biology*, 7(3), 903-911, available: http://dx.doi.org/10.1021/acssynbio.7b00405.

Chia, H.E., Marsh, E.N.G. and Biteen, J.S. (2019) 'Extending fluorescence microscopy into anaerobic environments', *Current Opinion in Chemical Biology*, 51, 98-104, available: http://dx.doi.org/10.1016/j.cbpa.2019.05.008.

Chiappero, M., Berruti, F., Mašek, O. and Fiore, S. (2021) 'Semi-continuous anaerobic digestion of mixed wastewater sludge with biochar addition', *Bioresource Technology*, 340, 125664, available: http://dx.doi.org/10.1016/j.biortech.2021.125664.

Chiarello, M., McCauley, M., Villéger, S. and Jackson, C.R. (2022) 'Ranking the biases: The choice of OTUs vs. ASVs in 16S rRNA amplicon data analysis has stronger effects on diversity measures than rarefaction and OTU identity threshold', *PLOS ONE*, 17(2), e0264443, available: http://dx.doi.org/10.1371/journal.pone.0264443.

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X. and Ruden, D.M. (2012) 'A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3', *Fly*, 6(2), 80-92, available: http://dx.doi.org/10.4161/fly.19695.

Cirne, D.G., Björnsson, L., Alves, M. and Mattiasson, B. (2006) 'Effects of bioaugmentation by an anaerobic lipolytic bacterium on anaerobic digestion of lipid-rich waste', *Journal of Chemical Technology & Biotechnology*, 81(11), 1745-1752, available: http://dx.doi.org/10.1002/jctb.1597.

Clark, J.H. (2019) 'Green biorefinery technologies based on waste biomass', *Green Chemistry*, 21(6), 1168-1170, available: http://dx.doi.org/10.1039/c9gc90021g.

Clarke, D.J. (2010) 'The Rcs phosphorelay: more than just a two-component pathway', *Future Microbiology*, 5(8), 1173-84, available: http://dx.doi.org/10.2217/fmb.10.83.

Cleveland, C.C. and Yavitt, J.B. (1998) 'Microbial consumption of atmospheric isoprene in a temperate forest soil', *Applied and Environmental Microbiology*, 64(1), 172-7, available: http://dx.doi.org/10.1128/aem.64.1.172-177.1998.

Close, D.M., Ripp, S. and Sayler, G.S. (2009) 'Reporter proteins in whole-cell optical bioreporter detection systems, biosensor integrations, and biosensing applications', *Sensors*, 9(11), 9147-9174, available: http://dx.doi.org/10.3390/s91109147.

Coker, J., Zhalnina, K., Marotz, C., Thiruppathy, D., Tjuanta, M., D'Elia, G., Hailu, R., Mahosky, T., Rowan, M., Northen, T.R. and Zengler, K. (2022) 'A Reproducible and Tunable Synthetic Soil Microbial Community Provides New Insights into Microbial Ecology', *mSystems*, 7(6), e0095122, available: http://dx.doi.org/10.1128/msystems.00951-22.

Collivignarelli, M.C., Canato, M., Abbà, A. and Carnevale Miino, M. (2019) 'Biosolids: What are the different types of reuse?', *Journal of Cleaner Production*, 238, 117844, available: http://dx.doi.org/10.1016/j.jclepro.2019.117844.

Conklin, A., Stensel, H.D. and Ferguson, J. (2006) 'Growth kinetics and competition between Methanosarcina and Methanosaeta in mesophilic anaerobic digestion', *Water Environment Research*, 78(5), 486-96, available: http://dx.doi.org/10.2175/106143006x95393.

Conway, T. and Cohen, P.S. (2015) 'Commensal and Pathogenic Escherichia coli Metabolism in the Gut', *Microbiology Spectrum*, 3(3), MBP-0006-2014, available: http://dx.doi.org/10.1128/microbiolspec.MBP-0006-2014.

Cristobal-Cueto, P., García-Quintanilla, A., Esteban, J. and García-Quintanilla, M. (2021) 'Phages in Food Industry Biocontrol and Bioremediation', *Antibiotics*, 10(7), 786, available: http://dx.doi.org/10.3390/antibiotics10070786.

Cross, K., Beckman, N., Jahnes, B. and Sabree, Z.L. (2025) 'Microbiome metabolic capacity is buffered against phylotype losses by functional redundancy', *Applied and Environmental Microbiology*, 91(2), e0236824, available: http://dx.doi.org/10.1128/aem.02368-24.

Cruz, I.A., Andrade, L.R.S., Bharagava, R.N., Nadda, A.K., Bilal, M., Figueiredo, R.T. and Ferreira, L.F.R. (2021) 'An overview of process monitoring for anaerobic digestion', *Biosystems Engineering*, 207, 106-119, available: http://dx.doi.org/10.1016/j.biosystemseng.2021.04.008.

Curry, K.D., Wang, Q., Nute, M.G., Tyshaieva, A., Reeves, E., Soriano, S., Wu, Q., Graeber, E., Finzer, P., Mendling, W., Savidge, T., Villapol, S., Dilthey, A. and Treangen, T.J. (2022) 'Emu: species-level microbial community profiling of full-length 16S rRNA Oxford Nanopore sequencing data', *Nature Methods*, 19(7), 845-853, available: http://dx.doi.org/10.1038/s41592-022-01520-4.

Da Re, S., Le Quéré, B., Ghigo, J.-M. and Beloin, C. (2007) 'Tight Modulation of Escherichia coli Bacterial Biofilm Formation through Controlled Expression of Adhesion Factors', *Applied and Environmental Microbiology*, 73(10), 3391-3403, available: http://dx.doi.org/10.1128/AEM.02625-06.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R. and Genomes Project Analysis, G. (2011) 'The variant call format and VCFtools', *Bioinformatics*, 27(15), 2156-2158, available: http://dx.doi.org/10.1093/bioinformatics/btr330.

Dar, R.A., Tsui, T.-H., Zhang, L., Smoliński, A., Jurišić, V., Tong, Y.W., Aggarangsi, P. and Liu, R. (2025) 'Viruses in anaerobic digestion systems: Diversity, role and future prospects', *Critical Reviews in Environmental Science and Technology*, 55(12), 861-884, available: http://dx.doi.org/10.1080/10643389.2025.2457980.

Dargode, P.S., More, P.P., Gore, S.S., Asodekar, B.R., Sharma, M.B. and Lali, A.M. (2022) 'Microbial consortia adaptation to substrate changes in anaerobic digestion', *Preparative Biochemistry & Biotechnology*, 52(8), 924-936, available: http://dx.doi.org/10.1080/10826068.2021.2009859.

de Visser, J.A.G.M. and Lenski, R.E. (2002) 'Long-term experimental evolution in Escherichia coli. XI. Rejection of non-transitive interactions as cause of declining rate of adaptation', *BMC Evolutionary Biology*, 2(1), 19, available: http://dx.doi.org/10.1186/1471-2148-2-19.

De Vrieze, J., Christiaens, M.E.R., Walraedt, D., Devooght, A., Ijaz, U.Z. and Boon, N. (2017) 'Microbial community redundancy in anaerobic digestion drives process recovery after salinity exposure', *Water Research*, 111, 109-117, available: http://dx.doi.org/10.1016/j.watres.2016.12.042.

De Vrieze, J., Hennebel, T., Boon, N. and Verstraete, W. (2012) 'Methanosarcina: The rediscovered methanogen for heavy duty biomethanation', *Bioresource Technology*, 112, 1-9, available: http://dx.doi.org/10.1016/j.biortech.2012.02.079.

De Vrieze, J., Pinto, A.J., Sloan, W.T. and Ijaz, U.Z. (2018) 'The active microbial community more accurately reflects the anaerobic digestion process: 16S rRNA (gene) sequencing as a predictive tool', *Microbiome*, 6(1), 63, available: http://dx.doi.org/10.1186/s40168-018-0449-9.

De Vrieze, J. and Verstraete, W. (2016) 'Perspectives for microbial community composition in anaerobic digestion: from abundance and activity to connectivity', *Environmental Microbiology*, 18(9), 2797-809, available: http://dx.doi.org/10.1111/1462-2920.13437.

Del Valle, I., Fulk, E.M., Kalvapalle, P., Silberg, J.J., Masiello, C.A. and Stadler, L.B. (2020a) 'Translating New Synthetic Biology Advances for Biosensing Into the Earth and Environmental Sciences', *Frontiers in Microbiology*, 11, 618373, available: http://dx.doi.org/10.3389/fmicb.2020.618373.

Del Valle, I., Webster, T.M., Cheng, H.-Y., Thies, J.E., Kessler, A., Miller, M.K., Ball, Z.T., MacKenzie, K.R., Masiello, C.A., Silberg, J.J. and Lehmann, J. (2020b) 'Soil organic matter attenuates the efficacy of flavonoid-based plant-microbe communication', *Science Advances*, 6(5), eaax8254, available: http://dx.doi.org/10.1126/sciadv.aax8254.

DelaVega-Quintero, J.C., Nuñez-Pérez, J., Lara-Fiallos, M., Barba, P., Burbano-García, J.L. and Espín-Valladares, R. (2025) 'Advances and Challenges in Anaerobic Digestion for Biogas Production: Policy, Technological, and Microbial Perspectives', *Processes*, 13(11), 3648, available: http://dx.doi.org/10.3390/pr13113648.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P. and Andersen, G.L. (2006) 'Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB', *Applied and Environmental Microbiology*, 72(7), 5069-72, available: http://dx.doi.org/10.1128/aem.03006-05.

Dimitriu, T., Szczelkun, M.D. and Westra, E.R. (2020) 'Evolutionary Ecology and Interplay of Prokaryotic Innate and Adaptive Immune Systems', *Current Biology*, 30(19), R1189-R1202, available: http://dx.doi.org/10.1016/j.cub.2020.08.028.

Dione, N., Khelaifia, S., La Scola, B., Lagier, J.C. and Raoult, D. (2016) 'A quasi-universal medium to break the aerobic/anaerobic bacterial culture dichotomy in clinical microbiology', *Clinical Microbiology and Infection*, 22(1), 53-58, available: http://dx.doi.org/10.1016/j.cmi.2015.10.032.

Dong, H. and Zhang, D. (2014) 'Current development in genetic engineering strategies of Bacillus species', *Microbial Cell Factories*, 13(1), 63, available: http://dx.doi.org/10.1186/1475-2859-13-63.

Dong, H., Zhang, Y., Dai, Z. and Li, Y. (2010) 'Engineering clostridium strain to accept unmethylated DNA', *PLOS ONE*, 5(2), e9038, available: http://dx.doi.org/10.1371/journal.pone.0009038.

Dueholm, M.K.D., Andersen, K.S., Korntved, A.-K.C., Rudkjøbing, V., Alves, M., Bajón-Fernández, Y., Batstone, D., Butler, C., Cruz, M.C., Davidsson, Å., Erijman, L., Holliger, C., Koch, K., Kreuzinger, N., Lee, C., Lyberatos, G., Mutnuri, S., O'Flaherty, V., Oleskowicz-Popiel, P., Pokorna, D., Rajal, V., Recktenwald, M., Rodríguez, J., Saikaly, P.E., Tooker, N., Vierheilig, J., De Vrieze, J., Wurzbacher, C. and Nielsen, P.H. (2024) 'MiDAS 5: Global diversity of bacteria and archaea in anaerobic digesters', *Nature Communications*, 15(1), 5361, available: http://dx.doi.org/10.1038/s41467-024-49641-y.

Duller, S., Vrbancic, S., Szydłowski, Ł., Mahnert, A., Blohs, M., Predl, M., Kumpitsch, C., Zrim, V., Högenauer, C., Kosciolek, T., Schmitz, R.A., Eberhard, A., Dragovan, M., Schmidberger, L., Zurabischvili, T., Weinberger, V., Moser, A.M., Kolb, D., Pernitsch, D., Mohammadzadeh, R., Kühnast, T., Rattei, T. and Moissl-Eichinger, C. (2024) 'Targeted isolation of Methanobrevibacter strains from fecal samples expands the cultivated human archaeome', *Nature Communications*, 15(1), 7593, available: http://dx.doi.org/10.1038/s41467-024-52037-7.

Dyksma, S., Jansen, L. and Gallert, C. (2020) 'Syntrophic acetate oxidation replaces acetoclastic methanogenesis during thermophilic digestion of biowaste', *Microbiome*, 8(1), 105, available: http://dx.doi.org/10.1186/s40168-020-00862-5.

Eisen, J.A. (2000) 'Horizontal gene transfer among microbial genomes: new insights from complete genome analysis', *Current Opinion in Genetics & Development*, 10(6), 606-11, available: http://dx.doi.org/10.1016/s0959-437x(00)00143-x.

El-Sayed, S.E., Abdelaziz, N.A., Alshahrani, M.Y., El-Housseiny, G.S. and Aboshanab, K.M. (2024) 'Biologically active metabolites of Alcaligenes faecalis: diversity, statistical optimization, and future perspectives', *Future Science OA*, 10(1), 2430452, available: http://dx.doi.org/10.1080/20565623.2024.2430452.

Elmore, J.R., Dexter, G.N., Baldino, H., Huenemann, J.D., Francis, R., Peabody, G.L., Martinez-Baird, J., Riley, L.A., Simmons, T., Coleman-Derr, D., Guss, A.M. and Egbert, R.G. (2023) 'High-throughput genetic engineering of nonmodel and undomesticated bacteria via iterative site-specific genome integration', *Science*

Advances, 9(10), eade1285, available: http://dx.doi.org/10.1126/sciadv.ade1285.

Elsayed, A., Laqa Kakar, F., Mustafa Abdelrahman, A., Ahmed, N., AlSayed, A., Sherif Zagloul, M., Muller, C., Bell, K.Y., Santoro, D., Norton, J., Marcus, A. and Elbeshbishy, E. (2024) 'Enhancing anaerobic digestion Efficiency: A comprehensive review on innovative intensification technologies', *Energy Conversion and Management*, 320, 118979, available: http://dx.doi.org/10.1016/j.enconman.2024.118979.

Environment Agency (2007) *The Microbiology of Drinking Water (2007) - Part 13 - The isolation and enumeration of aerobic spore-forming bacteria by membrane filtration*, Bristol: Environment Agency.

Errington, J. and Aart, L.T.V. (2020) 'Microbe Profile: Bacillus subtilis: model organism for cellular development, and industrial workhorse', *Microbiology*, 166(5), 425-427, available: http://dx.doi.org/10.1099/mic.0.000922.

Ewels, P., Magnusson, M., Lundin, S. and Käller, M. (2016) 'MultiQC: summarize analysis results for multiple tools and samples in a single report', *Bioinformatics*, 32(19), 3047-3048, available: http://dx.doi.org/10.1093/bioinformatics/btw354.

Fan, D., Dai, D.-J. and Wu, H.-S. (2013) 'Ethylene Formation by Catalytic Dehydration of Ethanol with Industrial Considerations', *Materials*, 6(1), 101-115, available: http://dx.doi.org/10.3390/ma6010101.

Favale, N., Costa, S., Summa, D., Sabbioni, S., Mamolini, E., Tamburini, E. and Scapoli, C. (2025) 'Comparison of microbiome community structure and dynamics during anaerobic digestion of different renewable solid wastes', *Current Research in Microbial Sciences*, 8, 100383, available: http://dx.doi.org/10.1016/j.crmicr.2025.100383.

Favate, J.S., Liang, S., Cope, A.L., Yadavalli, S.S. and Shah, P. (2022) 'The landscape of transcriptional and translational changes over 22 years of bacterial adaptation', *eLife*, 11, e81979, available: http://dx.doi.org/10.7554/eLife.81979.

Ferguson, R.M.W., Coulon, F. and Villa, R. (2016) 'Organic loading rate: A promising microbial management tool in anaerobic digestion', *Water Research*, 100, 348-356, available: http://dx.doi.org/10.1016/j.watres.2016.05.009.

Fernandez-Gonzalez, N., Braz, G.H.R., Regueiro, L., Lema, J.M. and Carballa, M. (2021) 'Microbial invasions in sludge anaerobic digesters', *Applied Microbiology and Biotechnology*, 105(1), 21-33, available: http://dx.doi.org/10.1007/s00253-020-11009-9.

Finn, T.J., Shewaramani, S., Leahy, S.C., Janssen, P.H. and Moon, C.D. (2017) 'Dynamics and genetic diversification of Escherichia coli during experimental adaptation to an anaerobic environment', *PeerJ*, 5, e3244, available: http://dx.doi.org/10.7717/peerj.3244.

Fotidis, I.A., Karakashev, D. and Angelidaki, I. (2013) 'Bioaugmentation with an acetate-oxidising consortium as a tool to tackle ammonia inhibition of anaerobic digestion', *Bioresource Technology*, 146, 57-62, available: http://dx.doi.org/10.1016/j.biortech.2013.07.041.

Foulquier, C., Huang, C.-N., Nguyen, N.-P.-T., Thiel, A., Wilding-Steel, T., Soula, J., Yoo, M., Ehrenreich, A., Meynial-Salles, I., Liebl, W. and Soucaille, P. (2019) 'An efficient method for markerless mutant generation by allelic exchange in Clostridium acetobutylicum and Clostridium saccharobutylicum using suicide vectors', *Biotechnology for Biofuels*, 12(1), 31, available: http://dx.doi.org/10.1186/s13068-019-1364-4.

Franchini, A.G., Ihssen, J. and Egli, T. (2015) 'Effect of Global Regulators RpoS and Cyclic-AMP/CRP on the Catabolome and Transcriptome of Escherichia coli K12 during Carbon- and Energy-Limited Growth', *PLOS ONE*, 10(7), e0133793, available: http://dx.doi.org/10.1371/journal.pone.0133793.

Francis, D.M. and Page, R. (2010) 'Strategies to optimize protein expression in E. coli', *Current Protocols in Protein Science*, 61(1), 5.24.1-5.24.29, available: http://dx.doi.org/10.1002/0471140864.ps0524s61.

Franke-Whittle, I.H., Walter, A., Ebner, C. and Insam, H. (2014) 'Investigation into the effect of high concentrations of volatile fatty acids in anaerobic digestion on methanogenic communities', *Waste Management*, 34(11), 2080-9, available: http://dx.doi.org/10.1016/j.wasman.2014.07.020.

Frey, J., Schneider, F., Huhn, T., Spiteller, D., Schink, B. and Schleheck, D. (2018) 'Two enzymes of the acetone degradation pathway of Desulfococcus biacutus: coenzyme B12-dependent 2-hydroxyisobutyryl-CoA mutase and 3-hydroxybutyryl-CoA dehydrogenase', *Environmental Microbiology Reports*, 10(3), 283-292, available: http://dx.doi.org/10.1111/1758-2229.12637.

Fuerte-Stone, J. and Mimee, M. (2026) 'Advances in biocontainment strategies of engineered microbes for use in humans', *Current Opinion in Microbiology*, 91, 102727, available: http://dx.doi.org/10.1016/j.mib.2026.102727.

Fulk, E.M., Gao, X., Lu, L.C., Redeker, K.R., Masiello, C.A. and Silberg, J.J. (2022) 'Nondestructive Chemical Sensing within Bulk Soil Using 1000 Biosensors Per Gram of Matrix', *ACS Synthetic Biology*, 11(7), 2372-2383, available: http://dx.doi.org/10.1021/acssynbio.2c00083.

Fulk, E.M., Huh, D., Atkinson, J.T., Lie, M., Masiello, C.A. and Silberg, J.J. (2020) 'A Split Methyl Halide Transferase AND Gate That Reports by Synthesizing an Indicator Gas', *ACS Synthetic Biology*, 9(11), 3104-3113, available: http://dx.doi.org/10.1021/acssynbio.0c00355.

Galagan, J.E., Nusbaum, C., Roy, A., Endrizzi, M.G., Macdonald, P., FitzHugh, W., Calvo, S., Engels, R., Smirnov, S., Atnoor, D., Brown, A., Allen, N., Naylor, J., Stange-Thomann, N., DeArellano, K., Johnson, R., Linton, L., McEwan, P., McKernan, K., Talamas, J., Tirrell, A., Ye, W., Zimmer, A., Barber, R.D., Cann, I., Graham, D.E., Grahame, D.A., Guss, A.M., Hedderich, R., Ingram-Smith, C., Kuettner, H.C., Krzycki, J.A., Leigh, J.A., Li, W., Liu, J., Mukhopadhyay, B., Reeve, J.N., Smith, K., Springer, T.A., Umayam, L.A., White, O., White, R.H., Conway de Macario, E., Ferry, J.G., Jarrell, K.F., Jing, H., Macario, A.J., Paulsen, I., Pritchett, M., Sowers, K.R., Swanson, R.V., Zinder, S.H., Lander, E., Metcalf, W.W. and Birren, B. (2002) 'The genome of M. acetivorans reveals extensive metabolic and physiological diversity', *Genome Research*, 12(4), 532-42, available: http://dx.doi.org/10.1101/gr.223902.

Gallagher, R.R., Patel, J.R., Interiano, A.L., Rovner, A.J. and Isaacs, F.J. (2015) 'Multilayered genetic safeguards limit growth of microorganisms to defined environments', *Nucleic Acids Research*, 43(3), 1945-54, available: http://dx.doi.org/10.1093/nar/gku1378.

Gandiglio, M., Lanzini, A., Soto, A., Leone, P. and Santarelli, M. (2017) 'Enhancing the Energy Efficiency of Wastewater Treatment Plants through Co-digestion and Fuel Cell Systems', *Frontiers in Environmental Science*, 5, 70, available: http://dx.doi.org/10.3389/fenvs.2017.00070.

Garcia, S.N., Osburn, B.I. and Jay-Russell, M.T. (2020) 'One Health for Food Safety, Food Security, and Sustainable Food Production', *Frontiers in Sustainable Food Systems*, 4, 1, available: http://dx.doi.org/10.3389/fsufs.2020.00001.

García-Bayona, L. and Comstock, L.E. (2019) 'Streamlined Genetic Manipulation of Diverse Bacteroides and Parabacteroides Isolates from the Human Gut

Microbiota', *mBio*, 10(4), e01762-19, available: http://dx.doi.org/10.1128/mBio.01762-19.

Garrison, E. and Marth, G. (2012) 'Haplotype-based variant detection from short-read sequencing', *arXiv*, 1207.

Gasanov, U., Hughes, D. and Hansbro, P.M. (2005) 'Methods for the isolation and identification of Listeria spp. and Listeria monocytogenes: a review', *FEMS Microbiology Reviews*, 29(5), 851-875, available: http://dx.doi.org/10.1016/j.femsre.2004.12.002.

Gauger, E.J., Leatham, M.P., Mercado-Lubo, R., Laux, D.C., Conway, T. and Cohen, P.S. (2007) 'Role of motility and the flhDC Operon in Escherichia coli MG1655 colonization of the mouse intestine', *Infection and Immunity*, 75(7), 3315-24, available: http://dx.doi.org/10.1128/iai.00052-07.

Gholizadeh, O., Ghaleh, H.E.G., Tat, M., Ranjbar, R. and Dorostkar, R. (2024) 'The potential use of bacteriophages as antibacterial agents against Klebsiella pneumoniae', *Virology Journal*, 21(1), 191, available: http://dx.doi.org/10.1186/s12985-024-02450-7.

Giangeri, G., Tsapekos, P., Gaspari, M., Ghofrani-Isfahani, P., Treu, L., Kougias, P., Campanaro, S. and Angelidaki, I. (2024) 'A bioaugmentation strategy to recover methane production under sulfate-stressed conditions: Highlights on targeted sulfate-reducing bacteria and DIET-related species', *Applied Energy*, 362, 122940, available: http://dx.doi.org/10.1016/j.apenergy.2024.122940.

Gonçalves, O.S., Creevey, C.J. and Santana, M.F. (2023) 'Designing a synthetic microbial community through genome metabolic modeling to enhance plant–microbe interaction', *Environmental Microbiome*, 18(1), 81, available: http://dx.doi.org/10.1186/s40793-023-00536-3.

Gorla, G., Iturrioz, J.A., Esbensen, K.H. and Amigo, J.M. (2026) 'Getting “sampling” right for reliable process monitoring: main challenges and solutions', *Frontiers in Analytical Science*, 6, 1839552, available: http://dx.doi.org/10.3389/frans.2026.1839552.

Goswami, R., Chattopadhyay, P., Shome, A., Banerjee, S.N., Chakraborty, A.K., Mathew, A.K. and Chaudhury, S. (2016) 'An overview of physico-chemical mechanisms of biogas production by microbial communities: a step towards sustainable waste management', *3 Biotech*, 6(1), 72, available: http://dx.doi.org/10.1007/s13205-016-0395-9.

Green, J., Stapleton, M.R., Smith, L.J., Artymiuk, P.J., Kahramanoglou, C., Hunt, D.M. and Buxton, R.S. (2014) 'Cyclic-AMP and bacterial cyclic-AMP receptor proteins revisited: adaptation for different ecological niches', *Current Opinion in Microbiology*, 18(100), 1-7, available: http://dx.doi.org/10.1016/j.mib.2014.01.003.

Grob, R.L. (1985) *Modern practice of gas chromatography*, 2nd ed., New York: Wiley.

Guo, H., Oosterkamp, M.J., Tonin, F., Hendriks, A., Nair, R., van Lier, J.B. and de Kreuk, M. (2021) 'Reconsidering hydrolysis kinetics for anaerobic digestion of waste activated sludge applying cascade reactors with ultra-short residence times', *Water Research*, 202, 117398, available: http://dx.doi.org/10.1016/j.watres.2021.117398.

Guo, S., Toth, C.R.A., Luo, F., Chen, X., Xiao, J. and Edwards, E.A. (2022) 'Transient Oxygen Exposure Causes Profound and Lasting Changes to a Benzene-Degrading Methanogenic Community', *Environmental Science & Technology*, 56(18), 13036-13045, available: http://dx.doi.org/10.1021/acs.est.2c02624.

Gustafsson, M., Meneghetti, R., Souza Marques, F., Trim, H., Dong, R., Al Seadi, T., Rasi, S., Thual, J., Kornatz, P., Wall, D., Berntsen, I.C., Saxegaard, S., Lyng, K.-A., Nägele, H.-J., Heaven, S. and Bywater, A. (2024) *A perspective on the state of the biogas industry in 12 member countries of IEA Bioenergy Task 37*: IEA Bioenergy.

Gyulev, I.S., Willson, B.J., Hennessy, R.C., Krabben, P., Jenkinson, E.R. and Thomas, G.H. (2018) 'Part by Part: Synthetic Biology Parts Used in Solventogenic Clostridia', *ACS Synthetic Biology*, 7(2), 311-327, available: http://dx.doi.org/10.1021/acssynbio.7b00327.

Haffiez, N., Azizi, S.M.M., Zakaria, B.S. and Dhar, B.R. (2022) 'Propagation of antibiotic resistance genes during anaerobic digestion of thermally hydrolyzed sludge and their correlation with extracellular polymeric substances', *Scientific Reports*, 12(1), 6749, available: http://dx.doi.org/10.1038/s41598-022-10764-1.

Hagen, L.H., Vivekanand, V., Pope, P.B., Eijsink, V.G. and Horn, S.J. (2015) 'The effect of storage conditions on microbial community composition and biomethane potential in a biogas starter culture', *Applied Microbiology and Biotechnology*, 99(13), 5749-61, available: http://dx.doi.org/10.1007/s00253-015-6623-0.

Hanreich, A., Schimpf, U., Zakrzewski, M., Schlüter, A., Benndorf, D., Heyer, R., Rapp, E., Pühler, A., Reichl, U. and Klocke, M. (2013) 'Metagenome and metaproteome

analyses of microbial communities in mesophilic biogas-producing anaerobic batch fermentations indicate concerted plant carbohydrate degradation', *Systematic and Applied Microbiology*, 36(5), 330-338, available: http://dx.doi.org/10.1016/j.syapm.2013.03.006.

Hare, K.M., Smith-Vaughan, H.C., Beissbarth, J. and Leach, A.J. (2015) 'Haemophilus influenzae isolates survive for up to 20 years at -70 °C in skim milk tryptone glucose glycerol broth (STGGB) if thawing is avoided during re-culture', *Journal of Microbiological Methods*, 119, 132-3, available: http://dx.doi.org/10.1016/j.mimet.2015.10.013.

Harirchi, S., Lak, M., Dalkilic, K., Ningsih, L.M., Ozturk, A.B., Akan, A.P., Canh, N.D., Taherzadeh, M.J. and Kumar, G. (2026) 'Impact of emerging contaminants on inhibited steady-state anaerobic digestion: a review of the fate and degradation characteristics of antibiotics, PFAS, and microplastics', *Reviews in Environmental Science and Bio/Technology*, 25(2), 32, available: http://dx.doi.org/10.1007/s11157-026-09779-1.

Harirchi, S., Wainaina, S., Sar, T., Nojoumi, S.A., Parchami, M., Varjani, S., Khanal, S.K., Wong, J., Awasthi, M.K. and Taherzadeh, M.J. (2022) 'Microbiological insights into anaerobic digestion for biogas, hydrogen or volatile fatty acids (VFAs): a review', *Bioengineered*, 13(3), 6521-6557, available: http://dx.doi.org/10.1080/21655979.2022.2035986.

Harley-Nyang, D., Memon, F.A., Jones, N. and Galloway, T. (2022) 'Investigation and analysis of microplastics in sewage sludge and biosolids: A case study from one wastewater treatment works in the UK', *Science of The Total Environment*, 823, 153735, available: http://dx.doi.org/10.1016/j.scitotenv.2022.153735.

Hartig, A.M., Dai, W., Zhang, K., Kapoor, K., Rottinghaus, A.G., Moon, T.S. and Parker, K.M. (2024) 'Influence of Environmental Conditions on the Escape Rates of Biocontained Genetically Engineered Microbes', *Environmental Science & Technology*, 58(51), 22657-22667, available: http://dx.doi.org/10.1021/acs.est.4c10893.

Hassa, J., Klang, J., Benndorf, D., Pohl, M., Hülsemann, B., Mächtig, T., Effenberger, M., Pühler, A., Schlüter, A. and Theuerl, S. (2021) 'Indicative Marker Microbiome Structures Deduced from the Taxonomic Inventory of 67 Full-Scale Anaerobic Digesters of 49 Agricultural Biogas Plants', *Microorganisms*, 9(7), 1457, available: http://dx.doi.org/10.3390/microorganisms9071457.

Hatzenpichler, R., Connon Stephanie, A., Goudeau, D., Malmstrom Rex, R., Woyke, T. and Orphan Victoria, J. (2016) 'Visualizing in situ translational activity for

identifying and sorting slow-growing archaeal–bacterial consortia', *Proceedings of the National Academy of Sciences of the United States of America*, 113(28), E4069-E4078, available: http://dx.doi.org/10.1073/pnas.1603757113.

Hibbing, M.E., Fuqua, C., Parsek, M.R. and Peterson, S.B. (2010) 'Bacterial competition: surviving and thriving in the microbial jungle', *Nature Reviews Microbiology*, 8(1), 15-25, available: http://dx.doi.org/10.1038/nrmicro2259.

Hickel, J., Brockway, P., Kallis, G., Keyßer, L., Lenzen, M., Slameršak, A., Steinberger, J. and Ürge-Vorsatz, D. (2021) 'Urgent need for post-growth climate mitigation scenarios', *Nature Energy*, 6(8), 766-768, available: http://dx.doi.org/10.1038/s41560-021-00884-9.

Hilton, S.K., Castro-Nallar, E., Pérez-Losada, M., Toma, I., McCaffrey, T.A., Hoffman, E.P., Siegel, M.O., Simon, G.L., Johnson, W.E. and Crandall, K.A. (2016) 'Metataxonomic and Metagenomic Approaches vs. Culture-Based Techniques for Clinical Pathology', *Frontiers in Microbiology*, 7, 484, available: http://dx.doi.org/10.3389/fmicb.2016.00484.

Hoornweg, D., Bhada-Tata, P. and Kennedy, C. (2013) 'Environment: Waste production must peak this century', *Nature*, 502(7473), 615-617, available: http://dx.doi.org/10.1038/502615a.

Howe, S., Liu, Z., Zuo, B. and Zhao, J. (2024) 'Culturomics: A critical approach in studying the roles of human and animal microbiota', *Animal Nutriomics*, 1, e6, available: http://dx.doi.org/10.1017/anr.2024.6.

Huang, Y., Sheth, R.U., Zhao, S., Cohen, L.A., Dabaghi, K., Moody, T., Sun, Y., Ricaurte, D., Richardson, M., Velez-Cortes, F., Blazejewski, T., Kaufman, A., Ronda, C. and Wang, H.H. (2023) 'High-throughput microbial culturomics using automation and machine learning', *Nature Biotechnology*, 41(10), 1424-1433, available: http://dx.doi.org/10.1038/s41587-023-01674-2.

Huening Katherine, A., Groves Joshua, T., Wildenthal John, A., Tabita, F.R. and North Justin, A. (2024) 'Escherichia coli possessing the dihydroxyacetone phosphate shunt utilize 5′-deoxynucleosides for growth', *Microbiology Spectrum*, 12(4), e03086-23, available: http://dx.doi.org/10.1128/spectrum.03086-23.

Humm, E., Chettiyar, R., Rubbi, L., Del Vecchio, G., Pellegrini, M., Ha Sung, M. and Gunsalus Robert, P. (2025) 'Complete genomes of the hydrogenotrophic methanogen species—Methanospirillum stamsii Pt1 and Methanospirillum

lacunae Ki8-1', *Microbiology Resource Announcements*, 14(10), e00736-25, available: http://dx.doi.org/10.1128/mra.00736-25.

Ibarra-Esparza, F.E., González-López, M.E., Ibarra-Esparza, J., Lara-Topete, G.O., Senés-Guerrero, C., Cansdale, A., Forrester, S., Chong, J.P.J. and Gradilla-Hernández, M.S. (2023) 'Implementation of anaerobic digestion for valorizing the organic fraction of municipal solid waste in developing countries: Technical insights from a systematic review', *Journal of Environmental Management*, 347, 118993, available: http://dx.doi.org/10.1016/j.jenvman.2023.118993.

Ibarra-Esparza, J., Ibarra-Esparza, F., González-López, M., García-González, A. and Gradilla-Hernández, M. (2025) 'Instrumentation and Continuous Monitoring for the Anaerobic Digestion Process: A Systematic Review', *IEEE Access*, 13, 193820-193837, available: http://dx.doi.org/10.1109/ACCESS.2025.3623647.

Innard, N. and Chong, J.P.J. (2022) 'The challenges of monitoring and manipulating anaerobic microbial communities', *Bioresource Technology*, 344, 126326, available: http://dx.doi.org/10.1016/j.biortech.2021.126326.

Isidro-Coxca, M.I., Ortiz-Jiménez, S. and Puente, J.L. (2024) 'Type 1 fimbria and P pili: regulatory mechanisms of the prototypical members of the chaperone-usher fimbrial family', *Archives of Microbiology*, 206(9), 373, available: http://dx.doi.org/10.1007/s00203-024-04092-3.

Issitt, T., Sweeney, S.T., Brackenbury, W.J. and Redeker, K.R. (2022) 'Sampling and Analysis of Low-Molecular-Weight Volatile Metabolites in Cellular Headspace and Mouse Breath', *Metabolites*, 12(7), 599, available: http://dx.doi.org/10.3390/metabo12070599.

Jain, C., Rodriguez, R.L., Phillippy, A.M., Konstantinidis, K.T. and Aluru, S. (2018) 'High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries', *Nature Communications*, 9(1), 5114, available: http://dx.doi.org/10.1038/s41467-018-07641-9.

James, J., Yarnall, B., Koranteng, A., Gibson, J., Rahman, T. and Doyle, D.A. (2021) 'Protein over-expression in Escherichia coli triggers adaptation analogous to antimicrobial resistance', *Microbial Cell Factories*, 20(1), 13, available: http://dx.doi.org/10.1186/s12934-020-01462-6.

Jassim, S.A., Limoges, R.G. and El-Cheikh, H. (2016) 'Bacteriophage biocontrol in wastewater treatment', *World Journal of Microbiology and Biotechnology*, 32(4), 70, available: http://dx.doi.org/10.1007/s11274-016-2028-1.

Jeong, K.J. (2025) 'Advancing microbial engineering through synthetic biology', *Journal of Microbiology*, 63(3), e2503100, available: http://dx.doi.org/10.71150/jm.2503100.

Jia, X., Bu, R., Zhao, T. and Wu, K. (2019) 'Sensitive and Specific Whole-Cell Biosensor for Arsenic Detection', *Applied and Environmental Microbiology*, 85(11), e00694-19, available: http://dx.doi.org/10.1128/AEM.00694-19.

Jiang, C., McIlroy, S.J., Qi, R., Petriglieri, F., Yashiro, E., Kondrotaite, Z. and Nielsen, P.H. (2021a) 'Identification of microorganisms responsible for foam formation in mesophilic anaerobic digesters treating surplus activated sludge', *Water Research*, 191, 116779, available: http://dx.doi.org/10.1016/j.watres.2020.116779.

Jiang, C., Peces, M., Andersen, M.H., Kucheryavskiy, S., Nierychlo, M., Yashiro, E., Andersen, K.S., Kirkegaard, R.H., Hao, L., Høgh, J., Hansen, A.A., Dueholm, M.S. and Nielsen, P.H. (2021b) 'Characterizing the growing microorganisms at species level in 46 anaerobic digesters at Danish wastewater treatment plants: A six-year survey on microbial community structure and key drivers', *Water Research*, 193, 116871, available: http://dx.doi.org/10.1016/j.watres.2021.116871.

Jin, X., Shen, H., Zhou, P., Yang, J., Yang, S., Ni, H., Yu, Y. and Zhang, Z. (2025) 'Research Progress on Sepsis Diagnosis and Monitoring Based on Omics Technologies: A Review', *Diagnostics*, 15(22), available: http://dx.doi.org/10.3390/diagnostics15222887.

Joseph, R.C., Kim, N.M. and Sandoval, N.R. (2018) 'Recent Developments of the Synthetic Biology Toolkit for Clostridium', *Frontiers in Microbiology*, 9, 154, available: http://dx.doi.org/10.3389/fmicb.2018.00154.

Jourdain, L. and Gu, W. (2025) 'Designing synthetic microbial communities for enhanced anaerobic waste treatment', *Applied and Environmental Microbiology*, 91(6), e00404-25, available: http://dx.doi.org/10.1128/aem.00404-25.

Kaijalainen, T. and Palmu, A. (2015) 'Long-term survival of Streptococcus pneumoniae, Haemophilus influenzae and Moraxella catarrhalis as isolates and in nasopharyngeal specimens in frozen STGG storage medium', *Journal of Microbiological Methods*, 114, 38-9, available: http://dx.doi.org/10.1016/j.mimet.2015.04.012.

Kalvapalle, P.B., Staubus, A., Dysart, M.J., Gambill, L., Reyes Gamas, K., Lu, L.C., Silberg, J.J., Stadler, L.B. and Chappell, J. (2025) 'Information storage across a microbial community using universal RNA barcoding', *Nature Biotechnology*, 44, 269-276, available: http://dx.doi.org/10.1038/s41587-025-02593-0.

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. and Ishiguro-Watanabe, M. (2023) 'KEGG for taxonomy-based analysis of pathways and genomes', *Nucleic Acids Research*, 51(D1), D587-D592, available: http://dx.doi.org/10.1093/nar/gkac963.

Kang, D.D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H. and Wang, Z. (2019) 'MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies', *PeerJ*, 7, e7359, available: http://dx.doi.org/10.7717/peerj.7359.

Kawasaki, K. and Kamagata, Y. (2017) 'Phosphate-Catalyzed Hydrogen Peroxide Formation from Agar, Gellan, and κ-Carrageenan and Recovery of Microbial Cultivability via Catalase and Pyruvate', *Applied and Environmental Microbiology*, 83(21), e01366-17, available: http://dx.doi.org/10.1128/aem.01366-17.

Keating, C., Trego, A., O'Flaherty, V. and Ijaz, U.Z. (2025) 'Microbiomes of high-rate anaerobic digestors reveal 'Study'-specific factors and limitations of synthetic wastewater', *Water Research*, 282, 123931, available: http://dx.doi.org/10.1016/j.watres.2025.123931.

Kerckhof, F.M., Courtens, E.N., Geirnaert, A., Hoefman, S., Ho, A., Vilchez-Vargas, R., Pieper, D.H., Jauregui, R., Vlaeminck, S.E., Van de Wiele, T., Vandamme, P., Heylen, K. and Boon, N. (2014) 'Optimized cryopreservation of mixed microbial communities for conserved functionality and diversity', *PLOS ONE*, 9(6), e99517, available: http://dx.doi.org/10.1371/journal.pone.0099517.

Kertesz, M.A. (2000) 'Riding the sulfur cycle – metabolism of sulfonates and sulfate esters in Gram-negative bacteria', *FEMS Microbiology Reviews*, 24(2), 135-175, available: http://dx.doi.org/10.1016/S0168-6445(99)00033-9.

Khan, M.A., Ngo, H.H., Guo, W.S., Liu, Y., Nghiem, L.D., Hai, F.I., Deng, L.J., Wang, J. and Wu, Y. (2016) 'Optimization of process parameters for production of volatile fatty acid, biohydrogen and methane from anaerobic digestion', *Bioresource Technology*, 219, 738-748, available: http://dx.doi.org/10.1016/j.biortech.2016.08.073.

Khanh Nguyen, V., Kumar Chaudhary, D., Hari Dahal, R., Hoang Trinh, N., Kim, J., Chang, S.W., Hong, Y., Duc La, D., Nguyen, X.C., Hao Ngo, H., Chung, W.J. and Nguyen, D.D. (2021) 'Review on pretreatment techniques to improve anaerobic digestion of sewage sludge', *Fuel*, 285, 119105, available: http://dx.doi.org/10.1016/j.fuel.2020.119105.

Kieft, B., Finke, N., McLaughlin, R.J., Nallan, A.N., Krzywinski, M., Crowe, S.A. and Hallam, S.J. (2023) 'Genome-resolved correlation mapping links microbial community structure to metabolic interactions driving methane production from wastewater', *Nature Communications*, 14(1), 5380, available: http://dx.doi.org/10.1038/s41467-023-40907-5.

Kim, J. and Kim, K.H. (2017) 'Effects of minimal media vs. complex media on the metabolite profiles of Escherichia coli and Saccharomyces cerevisiae', *Process Biochemistry*, 57, 64-71, available: http://dx.doi.org/10.1016/j.procbio.2017.04.003.

Kinsler, G., Schmidlin, K., Newell, D., Eder, R., Apodaca, S., Lam, G., Petrov, D. and Geiler-Samerotte, K. (2023) 'Extreme Sensitivity of Fitness to Environmental Conditions: Lessons from #1BigBatch', *Journal of Molecular Evolution*, 91(3), 293-310, available: http://dx.doi.org/10.1007/s00239-023-10114-3.

Kirby, M.E., Mirza, M.W., Hoskyns-Abrahall, H., Fenwick, J. and Theodorou, M.K. (2020) 'Increasing the Methane Potential of Oat Husks Using a Novel Extrusion Pre-Treatment Technology Prior to Anaerobic Digestion', *Frontiers in Energy Research*, 8, 579034, available: http://dx.doi.org/10.3389/fenrg.2020.579034.

Klein, W., Horlacher, R. and Boos, W. (1995) 'Molecular analysis of treB encoding the Escherichia coli enzyme II specific for trehalose', *Journal of Bacteriology*, 177(14), 4043-52, available: http://dx.doi.org/10.1128/jb.177.14.4043-4052.1995.

Klemm, P., Jørgensen, B.J., van Die, I., de Ree, H. and Bergmans, H. (1985) 'The fim genes responsible for synthesis of type 1 fimbriae in Escherichia coli, cloning and genetic organization', *Molecular and General Genetics*, 199(3), 410-4, available: http://dx.doi.org/10.1007/bf00330751.

Kolmogorov, M., Yuan, J., Lin, Y. and Pevzner, P.A. (2019) 'Assembly of long, error-prone reads using repeat graphs', *Nature Biotechnology*, 37(5), 540-546, available: http://dx.doi.org/10.1038/s41587-019-0072-8.

Kornienko, I.V., Aramova, O.Y., Tishchenko, A.A., Rudoy, D.V. and Chikindas, M.L. (2024) 'RNA Stability: A Review of the Role of Structural Features and Environmental Conditions', *Molecules*, 29(24), 5978, available: http://dx.doi.org/10.3390/molecules29245978.

Kougias, P.G., Boe, K., S, O.T., Kristensen, L.A. and Angelidaki, I. (2014) 'Anaerobic digestion foaming in full-scale biogas plants: a survey on causes and solutions', *Water Science and Technology*, 69(4), 889-95, available: http://dx.doi.org/10.2166/wst.2013.792.

Kulakowski, S., Banerjee, D., Scown, C.D. and Mukhopadhyay, A. (2023) 'Improving microbial bioproduction under low-oxygen conditions', *Current Opinion in Biotechnology*, 84, 103016, available: http://dx.doi.org/10.1016/j.copbio.2023.103016.

Kulasekara, H.D. and Blomfield, I.C. (1999) 'The molecular basis for the specificity of fimE in the phase variation of type 1 fimbriae of Escherichia coli K-12', *Molecular Microbiology*, 31(4), 1171-81, available: http://dx.doi.org/10.1046/j.1365-2958.1999.01257.x.

Laturnus, F., Haselmann, K.F., Borch, T. and Grøn, C. (2002) 'Terrestrial natural sources of trichloromethane (chloroform, CHCl3) – An overview', *Biogeochemistry*, 60(2), 121-139, available: http://dx.doi.org/10.1023/A:1019887505651.

Law, J.W.-F., Ab Mutalib, N.-S., Chan, K.-G. and Lee, L.-H. (2015) 'Rapid methods for the detection of foodborne bacterial pathogens: principles, applications, advantages and limitations', *Frontiers in Microbiology*, 5, 770, available: http://dx.doi.org/10.3389/fmicb.2014.00770.

Lawrence, D., Fiegna, F., Behrends, V., Bundy, J.G., Phillimore, A.B., Bell, T. and Barraclough, T.G. (2012) 'Species Interactions Alter Evolutionary Responses to a Novel Environment', *PLOS Biology*, 10(5), e1001330, available: http://dx.doi.org/10.1371/journal.pbio.1001330.

Lawson, C.E., Harcombe, W.R., Hatzenpichler, R., Lindemann, S.R., Löffler, F.E., O'Malley, M.A., García Martín, H., Pfleger, B.F., Raskin, L., Venturelli, O.S., Weissbrodt, D.G., Noguera, D.R. and McMahon, K.D. (2019) 'Common principles and best practices for engineering microbiomes', *Nature Reviews Microbiology*, 17(12), 725-741, available: http://dx.doi.org/10.1038/s41579-019-0255-9.

Leatham, M.P., Stevenson, S.J., Gauger, E.J., Krogfelt, K.A., Lins, J.J., Haddock, T.L., Autieri, S.M., Conway, T. and Cohen, P.S. (2005) 'Mouse intestine selects

nonmotile flhDC mutants of Escherichia coli MG1655 with increased colonizing ability and better utilization of carbon sources', *Infection and Immunity*, 73(12), 8039-49, available: http://dx.doi.org/10.1128/iai.73.12.8039-8049.2005.

Leca, E., Zennaro, B., Hamelin, J., Carrère, H. and Sambusiti, C. (2023) 'Use of additives to improve collective biogas plant performances: A comprehensive review', *Biotechnology Advances*, 65, 108129, available: http://dx.doi.org/10.1016/j.biotechadv.2023.108129.

Lee, B., Park, J.-G., Shin, W.-B., Tian, D.-J. and Jun, H.-B. (2017) 'Microbial communities change in an anaerobic digestion after application of microbial electrolysis cells', *Bioresource Technology*, 234, 273-280, available: http://dx.doi.org/10.1016/j.biortech.2017.02.022.

Lee, S.-H., Kang, H.-J., Lee, Y.H., Lee, T.J., Han, K., Choi, Y. and Park, H.-D. (2012) 'Monitoring bacterial community structure and variability in time scale in full-scale anaerobic digesters', *Journal of Environmental Monitoring*, 14(7), 1893-1905, available: http://dx.doi.org/10.1039/C2EM10958A.

Leite, G.B. and Hallenbeck, P.C. (2014) 'Engineered Cyanobacteria: Research and Application in Bioenergy' in Gupta, V. K., Tuohy, M. G., Kubicek, C. P., Saddler, J. and Xu, F., eds., *Bioenergy Research: Advances and Applications*, Amsterdam: Elsevier, 389-406.

Lenth, R. and Piaskowski, J. (2025) 'emmeans: Estimated Marginal Means, aka Least-Squares Means'.

Levantesi, C., Beimfohr, C., Blanch, A.R., Carducci, A., Gianico, A., Lucena, F., Tomei, M.C. and Mininni, G. (2015) 'Hygienization performances of innovative sludge treatment solutions to assure safe land spreading', *Environmental Science and Pollution Research*, 22(10), 7237-7247, available: http://dx.doi.org/10.1007/s11356-014-3572-6.

Li, C., Lü, F., Peng, W., He, P.-J. and Zhang, H. (2024) 'Functional Redundant Microbiome Enhanced Anaerobic Digestion Efficiency under Ammonium Inhibition Conditions', *Environmental Science & Technology*, 58(15), 6659-6669, available: http://dx.doi.org/10.1021/acs.est.4c01227.

Li, H. and Durbin, R. (2009) 'Fast and accurate short read alignment with Burrows-Wheeler transform', *Bioinformatics*, 25(14), 1754-60, available: http://dx.doi.org/10.1093/bioinformatics/btp324.

Li, L., Peng, X., Wang, X. and Wu, D. (2018) 'Anaerobic digestion of food waste: A review focusing on process stability', *Bioresource Technology*, 248, 20-28, available: http://dx.doi.org/10.1016/j.biortech.2017.07.012.

Li, M., Nian, R., Xian, M. and Zhang, H. (2018) 'Metabolic engineering for the production of isoprene and isopentenol by Escherichia coli', *Applied Microbiology and Biotechnology*, 102(18), 7725-7738, available: http://dx.doi.org/10.1007/s00253-018-9200-5.

Li, Y., Alaimo, C.P., Kim, M., Kado, N.Y., Peppers, J., Xue, J., Wan, C., Green, P.G., Zhang, R., Jenkins, B.M., Vogel, C.F.A., Wuertz, S., Young, T.M. and Kleeman, M.J. (2019) 'Composition and Toxicity of Biogas Produced from Different Feedstocks in California', *Environmental Science & Technology*, 53(19), 11569-11579, available: http://dx.doi.org/10.1021/acs.est.9b03003.

Lim, B., Zimmermann, M., Barry, N.A. and Goodman, A.L. (2017) 'Engineered Regulatory Systems Modulate Gene Expression of Human Commensals in the Gut', *Cell*, 169(3), 547-558.e15, available: http://dx.doi.org/10.1016/j.cell.2017.03.045.

Lim, J.W., Park, T., Tong, Y.W. and Yu, Z. (2020) 'The microbiome driving anaerobic digestion and microbial analysis', *Advances in Bioenergy*, 5, 1-61, available: http://dx.doi.org/10.1016/bs.aibe.2020.04.001.

Lima, D., Appleby, G. and Li, L. (2023) 'A Scoping Review of Options for Increasing Biogas Production from Sewage Sludge: Challenges and Opportunities for Enhancing Energy Self-Sufficiency in Wastewater Treatment Plants', *Energies*, 16(5), 2369, available: http://dx.doi.org/10.3390/en16052369.

Lindberg, P., Park, S. and Melis, A. (2010) 'Engineering a platform for photosynthetic isoprene production in cyanobacteria, using Synechocystis as the model organism', *Metabolic Engineering*, 12(1), 70-9, available: http://dx.doi.org/10.1016/j.ymben.2009.10.001.

Lindner, M.S. and Renard, B.Y. (2015) 'Metagenomic Profiling of Known and Unknown Microbes with MicrobeGPS', *PLOS ONE*, 10(2), e0117711, available: http://dx.doi.org/10.1371/journal.pone.0117711.

Liu, H. and Deutschbauer, A.M. (2018) 'Rapidly moving new bacteria to model-organism status', *Current Opinion in Biotechnology*, 51, 116-122, available: http://dx.doi.org/10.1016/j.copbio.2017.12.006.

Liu, L., Liu, Y., Shin, H.-D., Chen, R.R., Wang, N.S., Li, J., Du, G. and Chen, J. (2013) 'Developing Bacillus spp. as a cell factory for production of microbial enzymes and industrially important biochemicals in the context of systems and synthetic biology', *Applied Microbiology and Biotechnology*, 97(14), 6113-6127, available: http://dx.doi.org/10.1007/s00253-013-4960-4.

Liu, X., Pauss, A., André, L. and Ribeiro, T. (2025) 'An Upgraded FOS/TAC Titration Model Integrating Phosphate Effects for Accurate Assessments of Volatile Fatty Acids and Alkalinity in Anaerobic Media', *ChemEngineering*, 9(3), 53, available: http://dx.doi.org/10.3390/chemengineering9030053.

Liu, Y., Feng, J., Pan, H., Zhang, X. and Zhang, Y. (2022) 'Genetically engineered bacterium: Principles, practices, and prospects', *Frontiers in Microbiology*, 13, 997587, available: http://dx.doi.org/10.3389/fmicb.2022.997587.

Liu, Y., Guo, M., Du, R., Chi, J., He, X., Xie, Z., Huang, K., Luo, Y. and Xu, W. (2020) 'A gas reporting whole-cell microbial biosensor system for rapid on-site detection of mercury contamination in soils', *Biosensors and Bioelectronics*, 170, 112660, available: http://dx.doi.org/10.1016/j.bios.2020.112660.

Liu, Y.X., Qin, Y., Chen, T., Lu, M., Qian, X., Guo, X. and Bai, Y. (2021) 'A practical guide to amplicon and metagenomic analysis of microbiome data', *Protein Cell*, 12(5), 315-330, available: http://dx.doi.org/10.1007/s13238-020-00724-8.

Mackey, B.M., Kelly, A.F., Colvin, J.A., Robbins, P.T. and Fryer, P.J. (2006) 'Predicting the thermal inactivation of bacteria in a solid matrix: Simulation studies on the relative effects of microbial thermal resistance parameters and process conditions', *International Journal of Food Microbiology*, 107(3), 295-303, available: http://dx.doi.org/10.1016/j.ijfoodmicro.2005.10.008.

Madeira, F., Madhusoodanan, N., Lee, J., Eusebi, A., Niewielska, A., Tivey, A.R.N., Lopez, R. and Butcher, S. (2024) 'The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024', *Nucleic Acids Research*, 52(W1), W521-W525, available: http://dx.doi.org/10.1093/nar/gkae241.

Madigou, C., Poirier, S., Bureau, C. and Chapleur, O. (2016) 'Acclimation strategy to increase phenol tolerance of an anaerobic microbiota', *Bioresource Technology*, 216, 77-86, available: http://dx.doi.org/10.1016/j.biortech.2016.05.045.

Madsen, M., Holm-Nielsen, J.B. and Esbensen, K.H. (2011) 'Monitoring of anaerobic digestion processes: A review perspective', *Renewable and Sustainable Energy*

Reviews, 15(6), 3141-3155, available: http://dx.doi.org/10.1016/j.rser.2011.04.026.

Maki, J.J. and Looft, T. (2018) 'Megasphaera stantonii sp. nov., a butyrate-producing bacterium isolated from the cecum of a healthy chicken', *International Journal of Systematic and Evolutionary Microbiology*, 68(11), 3409-3415, available: http://dx.doi.org/10.1099/ijsem.0.002991.

Maklashina, E., Berthold Deborah, A. and Cecchini, G. (1998) 'Anaerobic Expression of Escherichia coli Succinate Dehydrogenase: Functional Replacement of Fumarate Reductase in the Respiratory Chain during Anaerobic Growth', *Journal of Bacteriology*, 180(22), 5989-5996, available: http://dx.doi.org/10.1128/jb.180.22.5989-5996.1998.

Mallawaarachchi, V., Wickramarachchi, A., Xue, H., Papudeshi, B., Grigson, S.R., Bouras, G., Prahl, R.E., Kaphle, A., Verich, A., Talamantes-Becerra, B., Dinsdale, E.A. and Edwards, R.A. (2024) 'Solving genomic puzzles: computational methods for metagenomic binning', *Briefings in Bioinformatics*, 25(5), bbae372, available: http://dx.doi.org/10.1093/bib/bbae372.

Martin, M. (2011) 'Cutadapt removes adapter sequences from high-throughput sequencing reads', *EMBnet.journal*, 17(1), 10-12, available: http://dx.doi.org/10.14806/ej.17.1.200.

Martínez-Carranza, E., Barajas, H., Alcaraz, L.D., Servín-González, L., Ponce-Soto, G.Y. and Soberón-Chávez, G. (2018) 'Variability of Bacterial Essential Genes Among Closely Related Bacteria: The Case of Escherichia coli', *Frontiers in Microbiology*, 9, 1059, available: http://dx.doi.org/10.3389/fmicb.2018.01059.

Martínez-García, E. and de Lorenzo, V. (2017) 'Molecular tools and emerging strategies for deep genetic/genomic refactoring of Pseudomonas', *Current Opinion in Biotechnology*, (47), 120-132, available: http://dx.doi.org/10.1016/j.copbio.2017.06.013.

Matsuo, Y., Komiya, S., Yasumizu, Y., Yasuoka, Y., Mizushima, K., Takagi, T., Kryukov, K., Fukuda, A., Morimoto, Y., Naito, Y., Okada, H., Bono, H., Nakagawa, S. and Hirota, K. (2021) 'Full-length 16S rRNA gene amplicon analysis of human gut microbiota using MinION™ nanopore sequencing confers species-level resolution', *BMC Microbiology*, 21(1), 35, available: http://dx.doi.org/10.1186/s12866-021-02094-5.

McAllister, K.N., Sorg, J.A. and Margolin, W. (2019) 'CRISPR Genome Editing Systems in the Genus Clostridium: a Timely Advancement', *Journal of Bacteriology*, 201(16), e00219-19, available: http://dx.doi.org/10.1128/JB.00219-19.

McDonald, M.J. (2019) 'Microbial Experimental Evolution – a proving ground for evolutionary theory and a tool for discovery', *EMBO reports*, 20(8), e46992, available: http://dx.doi.org/10.15252/embr.201846992.

McFarland, M. and Kaye, J. (1992) 'Chlorofluorocarbons and ozone', *Photochemistry and Photobiology*, 55(6), 911-29, available: http://dx.doi.org/10.1111/j.1751-1097.1992.tb08540.x.

Mee, M.T., Collins, J.J., Church, G.M. and Wang, H.H. (2014) 'Syntrophic exchange in synthetic microbial communities', *Proceedings of the National Academy of Sciences of the United States of America*, 111(20), E2149-E2156, available: http://dx.doi.org/10.1073/pnas.1405641111.

Meegoda, J.N., Li, B., Patel, K. and Wang, L.B. (2018) 'A Review of the Processes, Parameters, and Optimization of Anaerobic Digestion', *International Journal of Environmental Research and Public Health*, 15(10), 2224, available: http://dx.doi.org/10.3390/ijerph15102224.

Meier-Kolthoff, J.P., Hahnke, R.L., Petersen, J., Scheuner, C., Michael, V., Fiebig, A., Rohde, C., Rohde, M., Fartmann, B., Goodwin, L.A., Chertkov, O., Reddy, T.B.K., Pati, A., Ivanova, N.N., Markowitz, V., Kyrpides, N.C., Woyke, T., Göker, M. and Klenk, H.-P. (2014) 'Complete genome sequence of DSM 30083T, the type strain (U5/41T) of Escherichia coli, and a proposal for delineating subspecies in microbial taxonomy', *Standards in Genomic Sciences*, 9(1), 2, available: http://dx.doi.org/10.1186/1944-3277-9-2.

Meng, J., Young, G. and Chen, J. (2021) 'The Rcs System in Enterobacteriaceae: Envelope Stress Responses and Virulence Regulation', *Frontiers in Microbiology*, 12, 627104, available: http://dx.doi.org/10.3389/fmicb.2021.627104.

Mercado, J.V., Koyama, M. and Nakasaki, K. (2022) 'Short-term changes in the anaerobic digestion microbiome and biochemical pathways with changes in organic load', *Science of The Total Environment*, 813, 152585, available: http://dx.doi.org/10.1016/j.scitotenv.2021.152585.

Mesquita, D.P., Quintelas, C., Amaral, A.L. and Ferreira, E.C. (2017) 'Monitoring biological wastewater treatment processes: recent advances in spectroscopy

applications', *Reviews in Environmental Science and Bio/Technology*, 16(3), 395-424, available: http://dx.doi.org/10.1007/s11157-017-9439-9.

Mioduska, J., Grabowiec, A. and Hupka, J. (2023) 'Digestate Quality Originating from Kitchen Waste', *Applied Sciences*, 13(18), 10353, available: http://dx.doi.org/10.3390/app131810353.

Mirsoleimani Azizi, S.M., Zakaria, B.S. and Dhar, B.R. (2024) 'Low-temperature thermal hydrolysis for enhancing sludge anaerobic digestion and antibiotic resistance management: Significance of digester solids retention time', *Science of The Total Environment*, 917, 170392, available: http://dx.doi.org/10.1016/j.scitotenv.2024.170392.

Moeller, L. and Görsch, K. (2015) 'Foam formation in full-scale biogas plants processing biogenic waste', *Energy, Sustainability and Society*, 5(1), 1, available: http://dx.doi.org/10.1186/s13705-014-0031-7.

Moeller, L. and Zehnsdorf, A. (2016) 'Process upsets in a full-scale anaerobic digestion bioreactor: over-acidification and foam formation during biogas production', *Energy, Sustainability and Society*, 6(1), 30, available: http://dx.doi.org/10.1186/s13705-016-0095-7.

Morales-Polo, C., Cledera-Castro, M.D. and Moratilla Soria, B.Y. (2018) 'Reviewing the Anaerobic Digestion of Food Waste: From Waste Generation and Anaerobic Process to Its Perspectives', *Applied Sciences*, 8(10), 1804, available: http://dx.doi.org/10.3390/app8101804.

Moraskie, M., Roshid, M.H.O., O'Connor, G., Dikici, E., Zingg, J.-M., Deo, S. and Daunert, S. (2021) 'Microbial whole-cell biosensors: Current applications, challenges, and future perspectives', *Biosensors and Bioelectronics*, 191, 113359, available: http://dx.doi.org/10.1016/j.bios.2021.113359.

Morvan, C., Folgosa, F., Kint, N., Teixeira, M. and Martin-Verstraete, I. (2021) 'Responses of Clostridia to oxygen: from detoxification to adaptive strategies', *Environmental Microbiology*, 23(8), 4112-4125, available: http://dx.doi.org/10.1111/1462-2920.15665.

Mu, Z., Zou, Z., Yang, Y., Wang, W., Xu, Y., Huang, J., Cai, R., Liu, Y., Mo, Y., Wang, B., Dang, Y., Li, Y., Liu, Y., Jiang, Y., Tan, Q., Liu, X., Hu, C., Li, H., Wei, S., Lou, C., Yu, Y. and Wang, J. (2018) 'A genetically engineered Escherichia coli that senses and degrades tetracycline antibiotic residue', *Synthetic and Systems Biotechnology*, 3(3), 196-203, available: http://dx.doi.org/10.1016/j.synbio.2018.05.001.

Murphy, K.C. (2016) 'λ Recombination and Recombineering', *EcoSal Plus*, 7(1), ESP-0011-2015, available: http://dx.doi.org/10.1128/ecosalplus.ESP-0011-2015.

Murray, C.S., Gao, Y. and Wu, M. (2021) 'Re-evaluating the evidence for a universal genetic boundary among microbial species', *Nature Communications*, 12(1), 4059, available: http://dx.doi.org/10.1038/s41467-021-24128-2.

Myers, T. and Dykstra, C.M. (2024) 'Teaching old dogs new tricks: genetic engineering methanogens', *Applied and Environmental Microbiology*, 90(7), e0224723, available: http://dx.doi.org/10.1128/aem.02247-23.

Möller, K. and Müller, T. (2012) 'Effects of anaerobic digestion on digestate nutrient availability and crop growth: A review', *Engineering in Life Sciences*, 12(3), 242-257, available: http://dx.doi.org/10.1002/elsc.201100085.

Müller, C.M., Aberg, A., Straseviçiene, J., Emody, L., Uhlin, B.E. and Balsalobre, C. (2009) 'Type 1 fimbriae, a colonization factor of uropathogenic Escherichia coli, are controlled by the metabolic sensor CRP-cAMP', *PLOS Pathogens*, 5(2), e1000303, available: http://dx.doi.org/10.1371/journal.ppat.1000303.

Nakano, M.M. and Zuber, P. (1998) 'Anaerobic growth of a "strict aerobe" (Bacillus subtilis)', *Annual Review of Microbiology*, 52, 165-90, available: http://dx.doi.org/10.1146/annurev.micro.52.1.165.

Nakasaki, K., Nguyen, K.K., Ballesteros, F.C., Maekawa, T. and Koyama, M. (2020) 'Characterizing the microbial community involved in anaerobic digestion of lipid-rich wastewater to produce methane gas', *Anaerobe*, 61, 102082, available: http://dx.doi.org/10.1016/j.anaerobe.2019.102082.

Nemergut, D.R., Schmidt, S.K., Fukami, T., O'Neill, S.P., Bilinski, T.M., Stanish, L.F., Knelman, J.E., Darcy, J.L., Lynch, R.C., Wickey, P. and Ferrenberg, S. (2013) 'Patterns and processes of microbial community assembly', *Microbiology and Molecular Biology Reviews*, 77(3), 342-56, available: http://dx.doi.org/10.1128/mmbr.00051-12.

Nielsen, D.R., Leonard, E., Yoon, S.-H., Tseng, H.-C., Yuan, C. and Prather, K.L.J. (2009) 'Engineering alternative butanol production platforms in heterologous bacteria', *Metabolic Engineering*, 11(4), 262-273, available: http://dx.doi.org/10.1016/j.ymben.2009.05.003.

Nielsen, H.B. and Angelidaki, I. (2008) 'Codigestion of manure and industrial organic waste at centralized biogas plants: process imbalances and limitations', *Water Science and Technology*, 58(7), 1521-8, available: http://dx.doi.org/10.2166/wst.2008.507.

Nierychlo, M., Andersen, K.S., Xu, Y., Green, N., Jiang, C., Albertsen, M., Dueholm, M.S. and Nielsen, P.H. (2020) 'MiDAS 3: An ecosystem-specific reference database, taxonomy and knowledge platform for activated sludge and anaerobic digesters reveals species-level microbiome composition of activated sludge', *Water Research*, 182, 115955, available: http://dx.doi.org/10.1016/j.watres.2020.115955.

Nikoloudaki, O., Aheto, F., Di Cagno, R. and Gobbetti, M. (2024) 'Synthetic microbial communities: A gateway to understanding resistance, resilience, and functionality in spontaneously fermented food microbiomes', *Food Research International*, 192, 114780, available: http://dx.doi.org/10.1016/j.foodres.2024.114780.

Nkuna, R., Roopnarain, A., Rashama, C. and Adeleke, R. (2022) 'Insights into organic loading rates of anaerobic digestion for biogas production: a review', *Critical Reviews in Biotechnology*, 42(4), 487-507, available: http://dx.doi.org/10.1080/07388551.2021.1942778.

Nobu, M.K., Narihiro, T., Mei, R., Kamagata, Y., Lee, P.K.H., Lee, P.-H., McInerney, M.J. and Liu, W.-T. (2020) 'Catabolism and interactions of uncultured organisms shaped by eco-thermodynamics in methanogenic bioprocesses', *Microbiome*, 8(1), 111, available: http://dx.doi.org/10.1186/s40168-020-00885-y.

North, J.A., Miller, A.R., Wildenthal, J.A., Young, S.J. and Tabita, F.R. (2017) 'Microbial pathway for anaerobic 5'-methylthioadenosine metabolism coupled to ethylene formation', *Proceedings of the National Academy of Sciences of the United States of America*, 114(48), E10455-e10464, available: http://dx.doi.org/10.1073/pnas.1711625114.

North, J.A., Narrowe, A.B., Xiong, W., Byerly, K.M., Zhao, G., Young, S.J., Murali, S., Wildenthal, J.A., Cannon, W.R., Wrighton, K.C., Hettich, R.L. and Tabita, F.R. (2020a) 'A nitrogenase-like enzyme system catalyzes methionine, ethylene, and methane biogenesis', *Science*, 369(6507), 1094-1098, available: http://dx.doi.org/10.1126/science.abb6310.

North, J.A., Wildenthal, J.A., Erb, T.J., Evans, B.S., Byerly, K.M., Gerlt, J.A. and Tabita, F.R. (2020b) 'A bifunctional salvage pathway for two distinct S-adenosylmethionine by-products that is widespread in bacteria, including pathogenic Escherichia

coli', *Molecular Microbiology*, 113(5), 923-937, available: http://dx.doi.org/10.1111/mmi.14459.

Northen, T.R., Kleiner, M., Torres, M., Kovács, Á.T., Nicolaisen, M.H., Krzyżanowska, D.M., Sharma, S., Lund, G., Jelsbak, L., Baars, O., Kindtler, N.L., Wippel, K., Dinesen, C., Ferrarezi, J.A., Marian, M., Pioppi, A., Xu, X., Andersen, T., Geldner, N., Schulze-Lefert, P., Vorholt, J.A. and Garrido-Oter, R. (2024) 'Community standards and future opportunities for synthetic communities in plant–microbiota research', *Nature Microbiology*, 9(11), 2774-2784, available: http://dx.doi.org/10.1038/s41564-024-01833-4.

Noto Guillen, M., Rosener, B., Sayin, S. and Mitchell, A. (2021) 'Assembling stable syntrophic Escherichia coli communities by comprehensively identifying beneficiaries of secreted goods', *Cell Systems*, 12(11), 1064-1078.e7, available: http://dx.doi.org/10.1016/j.cels.2021.08.002.

Nurk, S., Meleshko, D., Korobeynikov, A. and Pevzner, P.A. (2017) 'metaSPAdes: a new versatile metagenomic assembler', *Genome Research*, 27(5), 824-834, available: http://dx.doi.org/10.1101/gr.213959.116.

Orth, J.D., Conrad, T.M., Na, J., Lerman, J.A., Nam, H., Feist, A.M. and Palsson, B. (2011) 'A comprehensive genome-scale reconstruction of Escherichia coli metabolism--2011', *Molecular Systems Biology*, 7, 535, available: http://dx.doi.org/10.1038/msb.2011.65.

Otte, J.M., Blackwell, N., Soos, V., Rughöft, S., Maisch, M., Kappler, A., Kleindienst, S. and Schmidt, C. (2018) 'Sterilization impacts on marine sediment---Are we able to inactivate microorganisms in environmental samples?', *FEMS Microbiology Ecology*, 94(12), fiy189, available: http://dx.doi.org/10.1093/femsec/fiy189.

Ou, B., Garcia, C., Wang, Y., Zhang, W. and Zhu, G. (2018) 'Techniques for chromosomal integration and expression optimization in Escherichia coli', *Biotechnology and Bioengineering*, 115(10), 2467-2478, available: http://dx.doi.org/10.1002/bit.26790.

Palù, M., Peprah, M., Tsapekos, P., Kougias, P., Campanaro, S., Angelidaki, I. and Treu, L. (2022) 'In-situ biogas upgrading assisted by bioaugmentation with hydrogenotrophic methanogens during mesophilic and thermophilic co-digestion', *Bioresource Technology*, 348, 126754, available: http://dx.doi.org/10.1016/j.biortech.2022.126754.

Park, H., Yeo, S., Ryu, C.B. and Huh, C.S. (2024) 'A streamlined culturomics case study for the human gut microbiota research', *Scientific Reports*, 14(1), 20361, available: http://dx.doi.org/10.1038/s41598-024-71370-x.

Park, S., Kim, C., Lee, D., Song, D.H., Cheon, K.C., Lee, H.S., Kim, S.J., Kim, J.C. and Lee, S.Y. (2017) 'Construction of Bacillus thuringiensis Simulant Strains Suitable for Environmental Release', *Applied and Environmental Microbiology*, 83(9), e00126-17, available: http://dx.doi.org/10.1128/aem.00126-17.

Park, S.-H., Ryu, S. and Kang, D.-H. (2012) 'Development of an Improved Selective and Differential Medium for Isolation of Salmonella spp', *Journal of Clinical Microbiology*, 50(10), 3222-3226, available: http://dx.doi.org/10.1128/jcm.01228-12.

Parker, D.J., Demetci, P. and Li, G.W. (2019) 'Rapid Accumulation of Motility-Activating Mutations in Resting Liquid Culture of Escherichia coli', *Journal of Bacteriology*, 201(19), e00259-19, available: http://dx.doi.org/10.1128/jb.00259-19.

Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.-A. and Hugenholtz, P. (2018) 'A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life', *Nature Biotechnology*, 36(10), 996-1004, available: http://dx.doi.org/10.1038/nbt.4229.

Patinios, C., de Vries, S.T., Diallo, M., Lanza, L., Verbrugge, P.L.J.V.Q., López-Contreras, A.M., van der Oost, J., Weusthuis, R.A. and Kengen, S.W.M. (2023) 'Multiplex genome engineering in Clostridium beijerinckii NCIMB 8052 using CRISPR-Cas12a', *Scientific Reports*, 13(1), 10153, available: http://dx.doi.org/10.1038/s41598-023-37220-y.

Patton, G.E., Meraz, J.C., Yin, M., Worthan, S.B., Williams, S. and Behringer, M.G. (2026) 'Historical contingency limits adaptive diversification in a spatially structured environment', *Evolution Letters*, 10(1), 118-134, available: http://dx.doi.org/10.1093/evlett/qraf048.

Peng, W., Beggio, G., Pivato, A., Zhang, H., Lü, F. and He, P. (2022) 'Applications of near infrared spectroscopy and hyperspectral imaging techniques in anaerobic digestion of bio-wastes: A review', *Renewable and Sustainable Energy Reviews*, 165, 112608, available: http://dx.doi.org/10.1016/j.rser.2022.112608.

Pennings, P.S., Ogbunugafor, C.B. and Hershberg, R. (2022) 'Reversion is most likely under high mutation supply when compensatory mutations do not fully restore

fitness costs', *G3: Genes, Genomes, Genetics*, 12(9), jkac190, available: http://dx.doi.org/10.1093/g3journal/jkac190.

Perez, M., Ntemiri, A., Tan, H., Harris, H.M.B., Roager, H.M., Ribière, C. and O'Toole, P.W. (2021) 'A synthetic consortium of 100 gut commensals modulates the composition and function in a colon model of the microbiome of elderly subjects', *Gut Microbes*, 13(1), 1-19, available: http://dx.doi.org/10.1080/19490976.2021.1919464.

Peterson, D., Bonham, K.S., Rowland, S., Pattanayak, C.W. and Klepac-Ceraj, V. (2021) 'Comparative Analysis of 16S rRNA Gene and Metagenome Sequencing in Pediatric Gut Microbiomes', *Frontiers in Microbiology*, 12, 670336, available: http://dx.doi.org/10.3389/fmicb.2021.670336.

Pilli, S., Yan, S., Tyagi, R.D. and Surampalli, R.Y. (2015) 'Thermal Pretreatment of Sewage Sludge to Enhance Anaerobic Digestion: A Review', *Critical Reviews in Environmental Science and Technology*, 45(6), 669-702, available: http://dx.doi.org/10.1080/10643389.2013.876527.

Piñeiro, M., García-Olmedo, F. and Diaz, I. (1994) 'Redox modulation of the expression of bacterial genes encoding cysteine-rich proteins in plant protoplasts', *Proceedings of the National Academy of Sciences of the United States of America*, 91(9), 3867-71, available: http://dx.doi.org/10.1073/pnas.91.9.3867.

Pontrelli, S., Chiu, T.-Y., Lan, E.I., Chen, F.Y.H., Chang, P. and Liao, J.C. (2018) 'Escherichia coli as a host for metabolic engineering', *Metabolic Engineering*, 50, 16-46, available: http://dx.doi.org/10.1016/j.ymben.2018.04.008.

Posit team (2025) 'RStudio: Integrated Development Environment for R'.

Prodan, A., Tremaroli, V., Brolin, H., Zwinderman, A.H., Nieuwdorp, M. and Levin, E. (2020) 'Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing', *PLOS ONE*, 15(1), e0227434, available: http://dx.doi.org/10.1371/journal.pone.0227434.

Puig-Castellví, F., Midoux, C., Guenne, A., Conteau, D., Franchi, O., Bureau, C., Madigou, C., Jouan-Rimbaud Bouveresse, D., Kroff, P., Mazéas, L., Rutledge, D.N., Gaval, G. and Chapleur, O. (2022) 'Metataxonomics, metagenomics and metabolomics analysis of the influence of temperature modification in full-scale anaerobic digesters', *Bioresource Technology*, 346, 126612, available: http://dx.doi.org/10.1016/j.biortech.2021.126612.

Puigbò, P., Guzmán, E., Romeu, A. and Garcia-Vallvé, S. (2007) 'OPTIMIZER: a web server for optimizing the codon usage of DNA sequences', *Nucleic Acids Research*, 35(Web Server Issue), W126-W131, available: http://dx.doi.org/10.1093/nar/gkm219.

Pyne, M.E., Bruder, M., Moo-Young, M., Chung, D.A. and Chou, C.P. (2014) 'Technical guide for genetic advancement of underdeveloped and intractable Clostridium', *Biotechnology Advances*, 32(3), 623-641, available: http://dx.doi.org/10.1016/j.biotechadv.2014.04.003.

Quince, C., Walker, A.W., Simpson, J.T., Loman, N.J. and Segata, N. (2017) 'Shotgun metagenomics, from sampling to analysis', *Nature Biotechnology*, 35(9), 833-844, available: http://dx.doi.org/10.1038/nbt.3935.

R Core Team (2025) 'R: A Language and Environment for Statistical Computing'.

Rabii, A., Aldin, S., Dahman, Y. and Elbeshbishy, E. (2019) 'A Review on Anaerobic Co-Digestion with a Focus on the Microbial Populations and the Effect of Multi-Stage Digester Configuration', *Energies*, 12(6), 1106, available: http://dx.doi.org/10.3390/en12061106.

Ranjan, R., Rani, A., Metwally, A., McGee, H.S. and Perkins, D.L. (2016) 'Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing', *Biochemical and Biophysical Research Communications*, 469(4), 967-77, available: http://dx.doi.org/10.1016/j.bbrc.2015.12.083.

Redeker, K.R., Davis, S. and Kalin, R.M. (2007) 'Isotope values of atmospheric halocarbons and hydrocarbons from Irish urban, rural, and marine locations', *Journal of Geophysical Research: Atmospheres*, 112(D16), D16307, available: http://dx.doi.org/10.1029/2006JD007784.

Redhead, S., Nieuwland, J., Esteves, S., Lee, D.H., Kim, D.W., Mathias, J., Cha, C.J., Toleman, M., Dinsdale, R., Guwy, A. and Hayhurst, E. (2020) 'Fate of antibiotic resistant E. coli and antibiotic resistance genes during full scale conventional and advanced anaerobic digestion of sewage sludge', *PLOS ONE*, 15(12), e0237283, available: http://dx.doi.org/10.1371/journal.pone.0237283.

Ren, J., Karna, S., Lee, H.-M., Yoo, S.M. and Na, D. (2019) 'Artificial transformation methodologies for improving the efficiency of plasmid DNA transformation and simplifying its use', *Applied Microbiology and Biotechnology*, 103(23), 9205-9215, available: http://dx.doi.org/10.1007/s00253-019-10173-x.

Renella, G. and Giagnoni, L. (2016) 'Light dazzles from the black box: whole-cell biosensors are ready to inform on fundamental soil biological processes', *Chemical and Biological Technologies in Agriculture*, 3(1), 8, available: http://dx.doi.org/10.1186/s40538-016-0059-3.

Reynolds, R., Hyun, S., Tully, B., Bien, J. and Levine, N.M. (2023) 'Identification of microbial metabolic functional guilds from large genomic datasets', *Frontiers in Microbiology*, 14, 1197329, available: http://dx.doi.org/10.3389/fmicb.2023.1197329.

Riley, L.A., Ji, L., Schmitz, R.J., Westpheling, J. and Guss, A.M. (2019) 'Rational development of transformation in Clostridium thermocellum ATCC 27405 via complete methylome analysis and evasion of native restriction-modification systems', *Journal of Industrial Microbiol and Biotechnology*, 46(9-10), 1435-1443, available: http://dx.doi.org/10.1007/s10295-019-02218-x.

Rockinger, U., Funk, M. and Winter, G. (2021) 'Current Approaches of Preservation of Cells During (freeze-) Drying', *Journal of Pharmaceutical Sciences*, 110(8), 2873-2893, available: http://dx.doi.org/10.1016/j.xphs.2021.04.018.

Rodríguez Amor, D. and Dal Bello, M. (2019) 'Bottom-Up Approaches to Synthetic Cooperation in Microbial Communities', *Life*, 9(1), 22, available: http://dx.doi.org/10.3390/life9010022.

Roell, G.W., Zha, J., Carr, R.R., Koffas, M.A., Fong, S.S. and Tang, Y.J. (2019) 'Engineering microbial consortia by division of labor', *Microbial Cell Factories*, 18(1), 35, available: http://dx.doi.org/10.1186/s12934-019-1083-3.

Ronda, C., Chen, S.P., Cabral, V., Yaung, S.J. and Wang, H.H. (2019) 'Metagenomic engineering of the mammalian gut microbiome in situ', *Nature Methods*, 16(2), 167-170, available: http://dx.doi.org/10.1038/s41592-018-0301-y.

Rossi, F., Amadoro, C. and Colavita, G. (2019) 'Members of the Lactobacillus Genus Complex (LGC) as Opportunistic Pathogens: A Review', *Microorganisms*, 7(5), 126, available: http://dx.doi.org/10.3390/microorganisms7050126.

Roy, B., Das, T. and Bhattacharyya, S. (2023) 'Overview on Old and New Biochemical Test for Bacterial Identification', *Journal of Surgical Case Reports and Images*, 1, 23-28, available: http://dx.doi.org/10.31579/2690-1897/163.

Rubin, B.E., Diamond, S., Cress, B.F., Crits-Christoph, A., Lou, Y.C., Borges, A.L., Shivram, H., He, C., Xu, M., Zhou, Z., Smith, S.J., Rovinsky, R., Smock, D.C.J., Tang, K., Owens, T.K., Krishnappa, N., Sachdeva, R., Barrangou, R., Deutschbauer, A.M., Banfield, J.F. and Doudna, J.A. (2022) 'Species- and site-specific genome editing in complex bacterial communities', *Nature Microbiology*, 7(1), 34-47, available: http://dx.doi.org/10.1038/s41564-021-01014-7.

Rutala, W.A. and Weber, D.J. (2015) 'Disinfection, Sterilization, and Control of Hospital Waste' in Bennett, J. E., Dolin, R. and Blaser, M. J., eds., *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases*, 8th ed., Philadelphia: Elsevier Saunders, 3294-3309.e4.

Saha, S., Basak, B., Hwang, J.-H., Salama, E.-S., Chatterjee, P.K. and Jeon, B.-H. (2020) 'Microbial Symbiosis: A Network towards Biomethanation', *Trends in Microbiology*, 28(12), 968-984, available: http://dx.doi.org/10.1016/j.tim.2020.03.012.

Saleski, T.E., Chung, M.T., Carruthers, D.N., Khasbaatar, A., Kurabayashi, K. and Lin, X.N. (2021) 'Optimized gene expression from bacterial chromosome by high-throughput integration and screening', *Science Advances*, 7(7), eabe1767, available: http://dx.doi.org/10.1126/sciadv.abe1767.

Sanders, E.R. (2012) 'Aseptic laboratory techniques: plating methods', *Journal of Visualized Experiments*, (63), e3064, available: http://dx.doi.org/10.3791/3064.

Santinello, D., Zampieri, G., Agostini, S., Müller, B., Favaro, L., Treu, L. and Campanaro, S. (2024) 'Process stability in anaerobic Digestion: Unveiling microbial signatures of full-scale reactor performance', *Chemical Engineering Journal*, 497, 154962, available: http://dx.doi.org/10.1016/j.cej.2024.154962.

Santos, A., van Aerle, R., Barrientos, L. and Martinez-Urtaza, J. (2020) 'Computational methods for 16S metabarcoding studies using Nanopore sequencing data', *Computational and Structural Biotechnology Journal*, 18, 296-305, available: http://dx.doi.org/10.1016/j.csbj.2020.01.005.

Santos, M.A., Carromeu-Santos, A., Quina, A.S., Antunes, M.A., Kristensen, T.N., Santos, M., Matos, M., Fragata, I. and Simões, P. (2024) 'Experimental Evolution in a Warming World: The Omics Era', *Molecular Biology and Evolution*, 41(8), msae148, available: http://dx.doi.org/10.1093/molbev/msae148.

Saïda, F., Uzan, M., Odaert, B. and Bontems, F. (2006) 'Expression of highly toxic genes in E. coli: special strategies and genetic tools', *Current Protein & Peptide Science*, 7(1), 47-56, available: http://dx.doi.org/10.2174/138920306775474095.

Schavemaker, P.E. and Lynch, M. (2022) 'Flagellar energy costs across the tree of life', *eLife*, 11, e77266, available: http://dx.doi.org/10.7554/eLife.77266.

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

Schäfer, H. and Eyice, Ö. (2019) 'Microbial Cycling of Methanethiol', *Current Issues in Molecular Biology*, 33, 173-182, available: http://dx.doi.org/10.21775/cimb.033.173.

Sczyrba, A., Hofmann, P., Belmann, P., Koslicki, D., Janssen, S., Dröge, J., Gregor, I., Majda, S., Fiedler, J., Dahms, E., Bremges, A., Fritz, A., Garrido-Oter, R., Jørgensen, T.S., Shapiro, N., Blood, P.D., Gurevich, A., Bai, Y., Turaev, D., DeMaere, M.Z., Chikhi, R., Nagarajan, N., Quince, C., Meyer, F., Balvočiūtė, M., Hansen, L.H., Sørensen, S.J., Chia, B.K.H., Denis, B., Froula, J.L., Wang, Z., Egan, R., Don Kang, D., Cook, J.J., Deltel, C., Beckstette, M., Lemaitre, C., Peterlongo, P., Rizk, G., Lavenier, D., Wu, Y.-W., Singer, S.W., Jain, C., Strous, M., Klingenberg, H., Meinicke, P., Barton, M.D., Lingner, T., Lin, H.-H., Liao, Y.-C., Silva, G.G.Z., Cuevas, D.A., Edwards, R.A., Saha, S., Piro, V.C., Renard, B.Y., Pop, M., Klenk, H.-P., Göker, M., Kyrpides, N.C., Woyke, T., Vorholt, J.A., Schulze-Lefert, P., Rubin, E.M., Darling, A.E., Rattei, T. and McHardy, A.C. (2017) 'Critical Assessment of Metagenome Interpretation—a benchmark of metagenomics software', *Nature Methods*, 14(11), 1063-1071, available: http://dx.doi.org/10.1038/nmeth.4458.

Senaratne, W. and Jayaweera, J. (2024) 'Comparison of microbial preservation methods: a narrative review', *Germs*, 14(4), 375-386, available: http://dx.doi.org/10.18683/germs.2024.1447.

Setubal, J.C. (2021) 'Metagenome-assembled genomes: concepts, analogies, and challenges', *Biophysical Reviews*, 13(6), 905-909, available: http://dx.doi.org/10.1007/s12551-021-00865-y.

Seuylemezian, A., Aronson, H.S., Tan, J., Lin, M., Schubert, W. and Vaishampayan, P. (2018) 'Development of a Custom MALDI-TOF MS Database for Species-Level Identification of Bacterial Isolates Collected From Spacecraft and Associated Surfaces', *Frontiers in Microbiology*, 9, 780, available: http://dx.doi.org/10.3389/fmicb.2018.00780.

Sezonov, G., Joseleau-Petit, D. and D'Ari, R. (2007) 'Escherichia coli physiology in Luria-Bertani broth', *Journal of Bacteriology*, 189(23), 8746-9, available: http://dx.doi.org/10.1128/jb.01368-07.

Shen, W., Le, S., Li, Y. and Hu, F. (2016) 'SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation', *PLOS ONE*, 11(10), e0163962, available: http://dx.doi.org/10.1371/journal.pone.0163962.

Shepherd, E.S., DeLoache, W.C., Pruss, K.M., Whitaker, W.R. and Sonnenburg, J.L. (2018) 'An exclusive metabolic niche enables strain engraftment in the gut microbiota', *Nature*, 557(7705), 434-438, available: http://dx.doi.org/10.1038/s41586-018-0092-4.

Sheth, R.U., Cabral, V., Chen, S.P. and Wang, H.H. (2016) 'Manipulating Bacterial Communities by in situ Microbiome Engineering', *Trends in Genetics*, 32(4), 189-200, available: http://dx.doi.org/10.1016/j.tig.2016.01.005.

Shimada, T., Fujita, N., Yamamoto, K. and Ishihama, A. (2011) 'Novel roles of cAMP receptor protein (CRP) in regulation of transport and metabolism of carbon sources', *PLOS ONE*, 6(6), e20081, available: http://dx.doi.org/10.1371/journal.pone.0020081.

Shin, S.G., Kim, S.I. and Hwang, S. (2022) 'Startup of Demo-Scale Anaerobic Digestion Plant Treating Food Waste Leachate: Process Instability and Recovery', *International Journal of Environmental Research and Public Health*, 19(11), 6903, available: http://dx.doi.org/10.3390/ijerph19116903.

Shinde, R., Hackula, A., O'Shea, R., Barth, S., Murphy, J.D. and Wall, D.M. (2023) 'Demand-driven biogas production from Upflow Anaerobic Sludge Blanket (UASB) reactors to balance the power grid', *Bioresource Technology*, 385, 129364, available: http://dx.doi.org/10.1016/j.biortech.2023.129364.

Sieber, C.M.K., Probst, A.J., Sharrar, A., Thomas, B.C., Hess, M., Tringe, S.G. and Banfield, J.F. (2018) 'Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy', *Nature Microbiology*, 3(7), 836-843, available: http://dx.doi.org/10.1038/s41564-018-0171-1.

Singh, H.B., O'Hara, D., Herlth, D., Sachse, W., Blake, D.R., Bradshaw, J.D., Kanakidou, M. and Crutzen, P.J. (1994) 'Acetone in the atmosphere: Distribution, sources, and sinks', *Journal of Geophysical Research: Atmospheres*, 99(D1), 1805-1819, available: http://dx.doi.org/10.1029/93JD00764.

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

Smith, R.H., Glendinning, L., Walker, A.W. and Watson, M. (2022) 'Investigating the impact of database choice on the accuracy of metagenomic read classification for the rumen microbiome', *Animal Microbiome*, 4(1), 57, available: http://dx.doi.org/10.1186/s42523-022-00207-7.

Snipen, L., Angell, I.L., Rognes, T. and Rudi, K. (2021) 'Reduced metagenome sequencing for strain-resolution taxonomic profiles', *Microbiome*, 9(1), 79, available: http://dx.doi.org/10.1186/s40168-021-01019-8.

Sousa, A.M., Machado, I., Nicolau, A. and Pereira, M.O. (2013) 'Improvements on colony morphology identification towards bacterial profiling', *Journal of Microbiological Methods*, 95(3), 327-35, available: http://dx.doi.org/10.1016/j.mimet.2013.09.020.

Soutourina, O., Kolb, A., Krin, E., Laurent-Winter, C., Rimsky, S., Danchin, A. and Bertin, P. (1999) 'Multiple control of flagellum biosynthesis in Escherichia coli: role of H-NS protein and the cyclic AMP-catabolite activator protein complex in transcription of the flhDC master operon', *Journal of Bacteriology*, 181(24), 7500-8, available: http://dx.doi.org/10.1128/jb.181.24.7500-7508.1999.

Spanjers, H. and van Lier, J.B. (2006) 'Instrumentation in anaerobic treatment--research and practice', *Water Science and Technology*, 53(4-5), 63-76, available: http://dx.doi.org/10.2166/wst.2006.111.

Srinivasan, S., Beamer, M.A., Fiedler, T.L., Austin, M.N., Sizova, M.V., Strenk, S.M., Agnew, K.J., Gowda, G.A.N., Raftery, D., Epstein, S.S., Fredricks, D.N. and Hillier, S.L. (2019) 'Megasphaera lornae sp. nov., Megasphaera hutchinsoni sp. nov., and Megasphaera vaginalis sp. nov.: novel bacteria isolated from the female genital tract', *International Journal of Systematic and Evolutionary Microbiology*, 71(3), 004702, available: http://dx.doi.org/10.1099/ijsem.0.004702.

Srinivasan, S.R., Gillies, A.T., Chang, L., Thompson, A.D. and Gestwicki, J.E. (2012) 'Molecular chaperones DnaK and DnaJ share predicted binding sites on most proteins in the E. coli proteome', *Molecular BioSystems*, 8(9), 2323-33, available: http://dx.doi.org/10.1039/c2mb25145k.

Steensels, J., Gallone, B., Voordeckers, K. and Verstrepen, K.J. (2019) 'Domestication of Industrial Microbes', *Current Biology*, 29(10), R381-R393, available: http://dx.doi.org/10.1016/j.cub.2019.04.025.

Stirling, F. and Silver, P.A. (2020) 'Controlling the Implementation of Transgenic Microbes: Are We Ready for What Synthetic Biology Has to Offer?', *Molecular Cell*, 78(4), 614-623, available: http://dx.doi.org/10.1016/j.molcel.2020.03.034.

Stocker, J., Balluch, D., Gsell, M., Harms, H., Feliciano, J., Daunert, S., Malik, K.A. and van der Meer, J.R. (2003) 'Development of a set of simple bacterial biosensors for quantitative and rapid measurements of arsenite and arsenate in potable water', *Environmental Science & Technology*, 37(20), 4743-50, available: http://dx.doi.org/10.1021/es034258b.

Streett, H.E., Kalis, K.M. and Papoutsakis, E.T. (2019) 'A Strongly Fluorescing Anaerobic Reporter and Protein-Tagging System for *Clostridium* Organisms Based on the Fluorescence-Activating and Absorption-Shifting Tag Protein (FAST)', *Applied and Environmental Microbiology*, 85(14), e00622-19, available: http://dx.doi.org/10.1128/AEM.00622-19.

Strube, M.L. (2021) 'RibDif: can individual species be differentiated by 16S sequencing?', *Bioinformatics Advances*, 1(1), vbab020, available: http://dx.doi.org/10.1093/bioadv/vbab020.

Studier, F.W. (2005) 'Protein production by auto-induction in high density shaking cultures', *Protein Expression and Purification*, 41(1), 207-34, available: http://dx.doi.org/10.1016/j.pep.2005.01.016.

Subbarao, P.M.V., D' Silva, T.C., Adlak, K., Kumar, S., Chandra, R. and Vijay, V.K. (2023) 'Anaerobic digestion as a sustainable technology for efficiently utilizing biomass in the context of carbon neutrality and circular economy', *Environmental Research*, 234, 116286, available: http://dx.doi.org/10.1016/j.envres.2023.116286.

Suga, M. (1958) 'Studies on Explosive Reaction of Ethylene Mixed with Oxygen or Air. I. Explosion Limit at Low Pressure', *Bulletin of the Chemical Society of Japan*, 31(5), 515-521, available: http://dx.doi.org/10.1246/bcsj.31.515.

Svennevik, O.K., Solheim, O.E., Beck, G., Sørland, G.H., Jonassen, K.R., Rus, E., Westereng, B., Horn, S.J., Higgins, M.J. and Nilsen, P.J. (2019) 'Effects of post anaerobic digestion thermal hydrolysis on dewaterability and moisture

distribution in digestates', *Water Science and Technology*, 80(7), 1338-1346, available: http://dx.doi.org/10.2166/wst.2019.379.

Swayambhu, G., Moscatello, N., Atilla-Gokcumen, G.E. and Pfeifer, B.A. (2020) 'Flux Balance Analysis for Media Optimization and Genetic Targets to Improve Heterologous Siderophore Production', *iScience*, 23(4), 101016, available: http://dx.doi.org/10.1016/j.isci.2020.101016.

Tale, V.P., Maki, J.S., Struble, C.A. and Zitomer, D.H. (2011) 'Methanogen community structure-activity relationship and bioaugmentation of overloaded anaerobic digesters', *Water Research*, 45(16), 5249-5256, available: http://dx.doi.org/10.1016/j.watres.2011.07.035.

Tanaka, T., Kawasaki, K., Daimon, S., Kitagawa, W., Yamamoto, K., Tamaki, H., Tanaka, M., Nakatsu, C.H. and Kamagata, Y. (2014) 'A hidden pitfall in the preparation of agar media undermines microorganism cultivability', *Applied and Environmental Microbiology*, 80(24), 7659-66, available: http://dx.doi.org/10.1128/aem.02741-14.

Tang, K.H.D. (2023) 'Bioaugmentation of anaerobic wastewater treatment sludge digestion: A perspective on microplastics removal', *Journal of Cleaner Production*, 387, 135864, available: http://dx.doi.org/10.1016/j.jclepro.2023.135864.

Tao, A., Pang, F., Huang, S., Yu, G., Li, B. and Wang, T. (2014) 'Characterisation of endophytic Bacillus thuringiensis strains isolated from wheat plants as biocontrol agents against wheat flag smut', *Biocontrol Science and Technology*, 24(8), 901-924, available: http://dx.doi.org/10.1080/09583157.2014.904502.

Taylor, G., Mordaka, P. and Heap, J. (2018) 'Start-Stop Assembly: A functionally scarless DNA assembly system optimized for metabolic engineering', *Nucleic Acids Research*, 47, available: http://dx.doi.org/10.1093/nar/gky1182.

Teeling, H. and Glöckner, F.O. (2012) 'Current opportunities and challenges in microbial metagenome analysis--a bioinformatic perspective', *Briefings in Bioinformatics*, 13(6), 728-42, available: http://dx.doi.org/10.1093/bib/bbs039.

The Standing Committee of Analysts (2016) *The Microbiology of Drinking Water (2016) - Part 4 - Methods for the isolation and enumeration of coliform bacteria and Escherichia coli (including E. coli O157:H7)*, Nottingham: Environment Agency.

The Standing Committee of Analysts (2021) *The Microbiology of Drinking Water (2021) - Part 6 - Methods for the isolation and enumeration of sulphite-reducing clostridia and Clostridium Perfringens by membrane filtration*, Nottingham: Environment Agency.

Theuerl, S., Klang, J. and Prochnow, A. (2019) 'Process Disturbances in Agricultural Biogas Production—Causes, Mechanisms and Effects on the Biogas Microbiome: A Review', *Energies*, 12(3), 365, available: http://dx.doi.org/10.3390/en12030365.

Thi Trang Nhung, T., Verma, S., Ponne, S., Meghwanshi, G.K., Schön, T. and Kumar, R. (2025) 'Bacteriophage-based strategies for biocontrol and treatment of infectious diseases', *Computational and Structural Biotechnology Journal*, 27, 2924-2932, available: http://dx.doi.org/10.1016/j.csbj.2025.06.046.

Tong, J., Fang, P., Zhang, J., Wei, Y., Su, Y. and Zhang, Y. (2019) 'Microbial community evolution and fate of antibiotic resistance genes during sludge treatment in two full-scale anaerobic digestion plants with thermal hydrolysis pretreatment', *Bioresource Technology*, 288, 121575, available: http://dx.doi.org/10.1016/j.biortech.2019.121575.

Toprak, E., Veres, A., Michel, J.B., Chait, R., Hartl, D.L. and Kishony, R. (2011) 'Evolutionary paths to antibiotic resistance under dynamically sustained drug selection', *Nature Genetics*, 44(1), 101-5, available: http://dx.doi.org/10.1038/ng.1034.

Tringe, S.G. and Hugenholtz, P. (2008) 'A renaissance for the pioneering 16S rRNA gene', *Current Opinion in Microbiology*, 11(5), 442-446, available: http://dx.doi.org/10.1016/j.mib.2008.09.011.

Tukey, J.W. (1949) 'Comparing Individual Means in the Analysis of Variance', *Biometrics*, 5(2), 99-114, available: http://dx.doi.org/10.2307/3001913.

Uddin, M.M. and Wright, M.M. (2023) 'Anaerobic digestion fundamentals, challenges, and technological advances', *Physical Sciences Reviews*, 8(9), 2819-2837, available: http://dx.doi.org/10.1515/psr-2021-0068.

Unden, G., Strecker, A., Kleefeld, A. and Kim, O.B. (2016) 'C4-Dicarboxylate Utilization in Aerobic and Anaerobic Growth', *EcoSal Plus*, 7(1), ESP-0021-2015, available: http://dx.doi.org/10.1128/ecosalplus.ESP-0021-2015.

Urtecho, G., Tripp, A.D., Insigne, K.D., Kim, H. and Kosuri, S. (2019) 'Systematic Dissection of Sequence Elements Controlling σ70 Promoters Using a Genomically Encoded Multiplexed Reporter Assay in Escherichia coli', *Biochemistry*, 58(11), 1539-1551, available: http://dx.doi.org/10.1021/acs.biochem.7b01069.

Van den Bergh, B., Swings, T., Fauvart, M. and Michiels, J. (2018) 'Experimental Design, Population Dynamics, and Diversity in Microbial Experimental Evolution', *Microbiology and Molecular Biology Reviews*, 82(3), e00008-18, available: http://dx.doi.org/10.1128/mmbr.00008-18.

van der Meer, J.R. and Belkin, S. (2010) 'Where microbiology meets microengineering: design and applications of reporter bacteria', *Nature Reviews Microbiology*, 8(7), 511-522, available: http://dx.doi.org/10.1038/nrmicro2392.

van der Woude, M.W. and Bäumler, A.J. (2004) 'Phase and antigenic variation in bacteria', *Clinical Microbiology Reviews*, 17(3), 581-611, available: http://dx.doi.org/10.1128/cmr.17.3.581-611.2004.

Van Landuyt, J., Oosterlinck, J. and De Vrieze, J. (2026) 'The anaerobic digestion microbiome is robust toward variation in the waste activated sludge feed', *ISME Communications*, 6(1), ycaf072, available: http://dx.doi.org/10.1093/ismeco/ycaf072.

Vanwonterghem, I., Jensen, P.D., Dennis, P.G., Hugenholtz, P., Rabaey, K. and Tyson, G.W. (2014a) 'Deterministic processes guide long-term synchronised population dynamics in replicate anaerobic digesters', *The ISME Journal*, 8(10), 2015-2028, available: http://dx.doi.org/10.1038/ismej.2014.50.

Vanwonterghem, I., Jensen, P.D., Ho, D.P., Batstone, D.J. and Tyson, G.W. (2014b) 'Linking microbial community structure, interactions and function in anaerobic digesters using new molecular techniques', *Current Opinion in Biotechnology*, 27, 55-64, available: http://dx.doi.org/10.1016/j.copbio.2013.11.004.

Vanwonterghem, I., Jensen, P.D., Rabaey, K. and Tyson, G.W. (2016) 'Genome-centric resolution of microbial diversity, metabolism and interactions in anaerobic digestion', *Environmental Microbiology*, 18(9), 3144-58, available: http://dx.doi.org/10.1111/1462-2920.13382.

Vellend, M. (2010) 'Conceptual synthesis in community ecology', *The Quarterly Review of Biology*, 85(2), 183-206, available: http://dx.doi.org/10.1086/652373.

Venkiteshwaran, K., Bocher, B., Maki, J. and Zitomer, D. (2015) 'Relating Anaerobic Digestion Microbial Community and Process Function', *Microbiol Insights*, 8(Suppl 2), 37-44, available: http://dx.doi.org/10.4137/mbi.s33593.

Verdecia, J., Vahdat, K. and Isache, C. (2019) 'Trivalvular infective endocarditis secondary to Granulicatella adiacens and Peptostreptococcus spp', *IDCases*, 17, e00545, available: http://dx.doi.org/10.1016/j.idcr.2019.e00545.

Wainaina, S., Lukitawesa, Kumar Awasthi, M. and Taherzadeh, M.J. (2019) 'Bioengineering of anaerobic digestion for volatile fatty acids, hydrogen or methane production: A critical review', *Bioengineered*, 10(1), 437-458, available: http://dx.doi.org/10.1080/21655979.2019.1673937.

Walker, M., Zhang, Y., Heaven, S. and Banks, C. (2009) 'Potential errors in the quantitative evaluation of biogas production in anaerobic digestion processes', *Bioresource Technology*, 100(24), 6339-6346, available: http://dx.doi.org/10.1016/j.biortech.2009.07.018.

Walter, A., Probst, M., Franke-Whittle, I.H., Ebner, C., Podmirseg, S.M., Etemadi-Shalamzari, M., Hupfauf, S. and Insam, H. (2019) 'Microbiota in anaerobic digestion of sewage sludge with and without co-substrates', *Water and Environment Journal*, 33(2), 214-222, available: http://dx.doi.org/10.1111/wej.12392.

Wang, D., Hunt, K.A., Abrahamson, B., Flinkstrom, Z., Tao, X., Tanner, R.S., DeLeón, K.B., Zhou, A., Zhou, J., McInerney, M.J., Winkler, M.-K.H., Stahl, D.A., Candry, P. and Pan, C. (2025) 'Higher-order microbial interactions revealed by comparative metabolic modeling of synthetic communities with varying species composition', *ISME Communications*, 5(1), ycaf142, available: http://dx.doi.org/10.1093/ismeco/ycaf142.

Wang, D., Hunt, K.A., Candry, P., Tao, X., Wofford, N.Q., Zhou, J., McInerney, M.J., Stahl, D.A., Tanner, R.S., Zhou, A., Winkler, M. and Pan, C. (2023) 'Cross-Feedings, Competition, and Positive and Negative Synergies in a Four-Species Synthetic Community for Anaerobic Degradation of Cellulose to Methane', *mBio*, 14(2), e0318922, available: http://dx.doi.org/10.1128/mbio.03189-22.

Wang, F., Yu, L., Ren, Y., Zhang, Q., He, S., Zhao, M., He, Z., Gao, Q. and Chen, J. (2024) 'An optimized culturomics strategy for isolation of human milk microbiota', *Frontiers in Microbiology*, 15, 1272062, available: http://dx.doi.org/10.3389/fmicb.2024.1272062.

Wang, G., Zhao, Z., Ke, J., Engel, Y., Shi, Y.-M., Robinson, D., Bingol, K., Zhang, Z., Bowen, B., Louie, K., Wang, B., Evans, R., Miyamoto, Y., Cheng, K., Kosina, S., De Raad, M., Silva, L., Luhrs, A., Lubbe, A., Hoyt, D.W., Francavilla, C., Otani, H., Deutsch, S., Washton, N.M., Rubin, E.M., Mouncey, N.J., Visel, A., Northen, T., Cheng, J.-F., Bode, H.B. and Yoshikuni, Y. (2019) 'CRAGE enables rapid activation of biosynthetic gene clusters in undomesticated bacteria', *Nature Microbiology*, 4(12), 2498-2510, available: http://dx.doi.org/10.1038/s41564-019-0573-8.

Wang, J., Sun, Y., Zhang, D., Broderick, T., Strawn, M., Santha, H., Pallansch, K., Deines, A. and Wang, Z.-W. (2022) 'Unblocking the rate-limiting step of the municipal sludge anaerobic digestion', *Water Environment Research*, 94(10), e10793, available: http://dx.doi.org/10.1002/wer.10793.

Wang, S., Fleming, R.T., Westbrook, E.M., Matsumura, P. and McKay, D.B. (2006) 'Structure of the Escherichia coli FlhDC Complex, a Prokaryotic Heteromeric Regulator of Transcription', *Journal of Molecular Biology*, 355(4), 798-808, available: http://dx.doi.org/10.1016/j.jmb.2005.11.020.

Wang, Y., Yang, J., He, J., Wen, X., Ji, W. and Wang, Y. (2023) 'Inhibition Effect of KHCO3 and KH2PO4 on Ethylene Explosion', *ACS Omega*, 8(8), 7566-7574, available: http://dx.doi.org/10.1021/acsomega.2c06894.

Wang, Z., Hu, Y., Wang, S., Wu, G. and Zhan, X. (2023) 'A critical review on dry anaerobic digestion of organic waste: Characteristics, operational conditions, and improvement strategies', *Renewable and Sustainable Energy Reviews*, 176, 113208, available: http://dx.doi.org/10.1016/j.rser.2023.113208.

Westerholm, M., Levén, L. and Schnürer, A. (2012) 'Bioaugmentation of syntrophic acetate-oxidizing culture in biogas reactors exposed to increasing levels of ammonia', *Applied and Environmental Microbiology*, 78(21), 7619-25, available: http://dx.doi.org/10.1128/aem.01637-12.

Westlund, Å.D., Hagland, E. and Rothman, M. (1998) 'Operational aspects on foaming in digesters caused by microthrix parvicella', *Water Science and Technology*, 38(8-9), 29-34, available: http://dx.doi.org/10.2166/wst.1998.0787.

Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*, Springer-Verlag New York.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S.,

Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V. and Yutani, H. (2019) 'Welcome to the Tidyverse', *Journal of Open Source Software*, 4, 1686, available: http://dx.doi.org/10.21105/joss.01686.

Wickham, H., François, R., Henry, L., Müller, K. and D, V. (2023) 'dplyr: A Grammar of Data Manipulation'.

Withey, S., Cartmell, E., Avery, L.M. and Stephenson, T. (2005) 'Bacteriophages—potential for application in wastewater treatment processes', *Science of The Total Environment*, 339(1), 1-18, available: http://dx.doi.org/10.1016/j.scitotenv.2004.09.021.

Wood, D.E., Lu, J. and Langmead, B. (2019) 'Improved metagenomic analysis with Kraken 2', *Genome Biology*, 20(1), 257, available: http://dx.doi.org/10.1186/s13059-019-1891-0.

Woods, C., Humphreys, C.M., Rodrigues, R.M., Ingle, P., Rowe, P., Henstra, A.M., Köpke, M., Simpson, S.D., Winzer, K. and Minton, N.P. (2019) 'A novel conjugal donor strain for improved DNA transfer into Clostridium spp', *Anaerobe*, 59, 184-191, available: http://dx.doi.org/10.1016/j.anaerobe.2019.06.020.

Wu, D., Li, L., Zhao, X., Peng, Y., Yang, P. and Peng, X. (2019) 'Anaerobic digestion: A review on process monitoring', *Renewable and Sustainable Energy Reviews*, 103, 1-12, available: http://dx.doi.org/10.1016/j.rser.2018.12.039.

Wu, Y.-W., Simmons, B.A. and Singer, S.W. (2016) 'MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets', *Bioinformatics*, 32(4), 605-607, available: http://dx.doi.org/10.1093/bioinformatics/btv638.

Xu, R.Z., Fang, S., Zhang, L., Huang, W., Shao, Q., Fang, F., Feng, Q., Cao, J. and Luo, J. (2021) 'Distribution patterns of functional microbial community in anaerobic digesters under different operational circumstances: A review', *Bioresource Technology*, 341, 125823, available: http://dx.doi.org/10.1016/j.biortech.2021.125823.

Xue, G.-P., Johnson, J.S. and Dalrymple, B.P. (1999) 'High osmolarity improves the electro-transformation efficiency of the gram-positive bacteria Bacillus subtilis and Bacillus licheniformis', *Journal of Microbiological Methods*, 34(3), 183-191, available: http://dx.doi.org/10.1016/S0167-7012(98)00087-6.

Yachi, S. and Loreau, M. (1999) 'Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis', *Proceedings of the National Academy of Sciences of the United States of America*, 96(4), 1463-1468, available: http://dx.doi.org/10.1073/pnas.96.4.1463.

Yan, Q. and Fong, S.S. (2017) 'Challenges and Advances for Genetic Engineering of Non-model Bacteria and Uses in Consolidated Bioprocessing', *Frontiers in Microbiology*, 8, 2060, available: http://dx.doi.org/10.3389/fmicb.2017.02060.

Yang, G., Han, L., Wen, J. and Zhou, S. (2013) 'Pseudomonas guangdongensis sp. nov., isolated from an electroactive biofilm, and emended description of the genus Pseudomonas Migula 1894', *International Journal of Systematic and Evolutionary Microbiology*, 63(Pt 12), 4599-4605, available: http://dx.doi.org/10.1099/ijs.0.054676-0.

Yang, H., Gao, X., Jin, Z.-C., Zhang, R., Ning, B. and Yan, X. (2025) 'Simple and Versatile Toolkit for Genetic Manipulation of Bacillus licheniformis', *ACS Synthetic Biology*, 15(1), 262-270, available: http://dx.doi.org/10.1021/acssynbio.5c00699.

Yang, M., Vander Elst, M., Smets, I., Zhang, H., Li, S., Baeyens, J. and Deng, Y. (2024) 'Reviewing Improved Anaerobic Digestion by Combined Pre-Treatment of Waste-Activated Sludge (WAS)', *Sustainability*, 16(15), 6419, available: http://dx.doi.org/10.3390/su16156419.

Yang, P., Condrich, A., Lu, L., Scranton, S., Hebner, C., Sheykhhasan, M. and Ali, M.A. (2024) 'Genetic Engineering in Bacteria, Fungi, and Oomycetes, Taking Advantage of CRISPR', *DNA*, 4(4), 427-454, available: http://dx.doi.org/10.3390/dna4040030.

Yang, S., Svoronos, S.A. and Pullammanappallil, P. (2022) 'Development of Inexpensive, Automatic, Real-Time Measurement System for On-Line Methane Content and Biogas Flowrate', *Waste and Biomass Valorization*, 13(12), 4839-4849, available: http://dx.doi.org/10.1007/s12649-022-01835-5.

Yap, M., O'Sullivan, O., O'Toole, P.W. and Cotter, P.D. (2022) 'Development of sequencing-based methodologies to distinguish viable from non-viable cells in a bovine milk matrix: A pilot study', *Frontiers in Microbiology*, 13, 1036643, available: http://dx.doi.org/10.3389/fmicb.2022.1036643.

Yarberry, A., Lansing, S., Luckarift, H., Diltz, R., Mulbry, W. and Yarwood, S. (2019) 'Effect of anaerobic digester inoculum preservation via lyophilization on methane

recovery', *Waste Management*, 87, 62-70, available: http://dx.doi.org/10.1016/j.wasman.2019.01.033.

Yin, Q., Sun, Y., Li, B., Feng, Z. and Wu, G. (2022) 'The r/K selection theory and its application in biological wastewater treatment processes', *Science of The Total Environment*, 824, 153836, available: http://dx.doi.org/10.1016/j.scitotenv.2022.153836.

Yip, A., McArthur, O.D., Ho, K.C., Aucoin, M.G. and Ingalls, B.P. (2024) 'Degradation of polyethylene terephthalate (PET) plastics by wastewater bacteria engineered via conjugation', *Microbial Biotechnology*, 17(9), e70015, available: http://dx.doi.org/10.1111/1751-7915.70015.

Yue, S.-J., Huang, P., Wang, W., Hu, H.-B. and Zhang, X.-H. (2024) 'Development of a whole-cell biosensor for detection of pyocyanin based on a transcriptional regulation factor', *Sensors and Actuators B: Chemical*, 403, 135132, available: http://dx.doi.org/10.1016/j.snb.2023.135132.

Yutin, N. and Galperin, M.Y. (2013) 'A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia', *Environmental Microbiology*, 15(10), 2631-41, available: http://dx.doi.org/10.1111/1462-2920.12173.

Zahan, Z., Othman, M.Z. and Rajendram, W. (2016) 'Anaerobic Codigestion of Municipal Wastewater Treatment Plant Sludge with Food Waste: A Case Study', *BioMed Research International*, 2016, 8462928, available: http://dx.doi.org/10.1155/2016/8462928.

Zakrzewski, M., Goesmann, A., Jaenicke, S., Jünemann, S., Eikmeyer, F., Szczepanowski, R., Al-Soud, W.A., Sørensen, S., Pühler, A. and Schlüter, A. (2012) 'Profiling of the metabolically active community from a production-scale biogas plant by means of high-throughput metatranscriptome sequencing', *Journal of Biotechnology*, 158(4), 248-258, available: http://dx.doi.org/10.1016/j.jbiotec.2012.01.020.

Zaks, D.P., Winchester, N., Kucharik, C.J., Barford, C.C., Paltsev, S. and Reilly, J.M. (2011) 'Contribution of anaerobic digesters to emissions mitigation and electricity generation under U.S. climate policy', *Environmental Science & Technology*, 45(16), 6735-42, available: http://dx.doi.org/10.1021/es104227y.

Zelazny, A. and Bibi, E. (1996) 'Biogenesis and Topology of Integral Membrane Proteins: Characterization of Lactose Permease−Chloramphenicol Acetyltransferase

Hybrids', *Biochemistry*, 35(33), 10872-10878, available: http://dx.doi.org/10.1021/bi960815d.

Zhang, J., Gao, Q., Zhang, Q., Wang, T., Yue, H., Wu, L., Shi, J., Qin, Z., Zhou, J., Zuo, J. and Yang, Y. (2017) 'Bacteriophage-prokaryote dynamics and interaction within anaerobic digestion processes across time and space', *Microbiome*, 5(1), 57, available: http://dx.doi.org/10.1186/s40168-017-0272-8.

Zhang, L., Gong, X., Wang, L., Guo, K., Cao, S. and Zhou, Y. (2021) 'Metagenomic insights into the effect of thermal hydrolysis pre-treatment on microbial community of an anaerobic digestion system', *Science of the Total Environment*, 791, 148096, available: http://dx.doi.org/10.1016/j.scitotenv.2021.148096.

Zhang, X., Wang, Y., Jiao, P., Zhang, M., Deng, Y., Jiang, C., Liu, X.-W., Lou, L., Li, Y., Zhang, X.-X. and Ma, L. (2024) 'Microbiome-functionality in anaerobic digesters: A critical review', *Water Research*, 249, 120891, available: http://dx.doi.org/10.1016/j.watres.2023.120891.

Zhang, Y., Jing, M., Lyu, L., Nie, L., Xu, X., Sun, R., Chen, S., He, S., Huang, P., Luo, W., Liang, J., Gao, G., Fan, K., Yang, T., Zhang, L., Fu, X., Allard, S.M., Gilbert, J.A., Zhang, J. and Chu, H. (2026) 'Principles for Rigorous Design and Application of Synthetic Microbial Communities', *Advanced Science*, 13(10), e14750, available: http://dx.doi.org/10.1002/advs.202514750.

Zhou, H., Brown, R.C. and Wen, Z. (2019) 'Anaerobic digestion of aqueous phase from pyrolysis of biomass: Reducing toxicity and improving microbial tolerance', *Bioresource Technology*, 292, 121976, available: http://dx.doi.org/10.1016/j.biortech.2019.121976.

Zhou, X., Wang, J., Liu, F., Liang, J., Zhao, P., Tsui, C.K.M. and Cai, L. (2022) 'Cross-kingdom synthetic microbiota supports tomato suppression of Fusarium wilt disease', *Nature Communications*, 13(1), 7890, available: http://dx.doi.org/10.1038/s41467-022-35452-6.