Identifying Metabolically Active Microbial Populations in Anaerobic Digestion Through Bioorthogonal Non-Canonical Amino Acid Tagging and Affinity-Based Cell Separation

Yumechris Amekan

PhD

University of York

Biology

December 2021

Abstract

Our understanding of microbial communities associated with anaerobic digestion (AD) currently relies strongly on metagenomic sequencing, which can reveal phylogenetic diversity, but does not provide information concerning microbial activity or the close associations that may form between syntrophic organisms. Approaches that facilitate charting of process-targeted variation in microbial community activities are important for understanding how the microbiology of AD functions as a single biological machine. Here I identify subsets of metabolically specialised microbes as they respond to substrate availability in AD systems using bioorthogonal non-canonical amino acid tagging (BONCAT) and affinity-based cell separation (ABCS). The results demonstrate the specific labelling, visualisation and separation of microbes that actively participate in volatile fatty acid (VFA) degradation and suggest a differential response to octanoic acid by specialists within the microbial community. Analysis of metagenomic sequences from a time series of BONCAT-ABCS samples reveals that this method enriched a distinct microbial community with genetically-derived metabolisms consistent with and changing according to the observed metabolic outcomes. Proteomic data generated from BONCAT-ABCS help to resolve the functional landscape of enriched microbial community and give better resolution of identified translationally/metabolically active taxa. This enrichment approach allows us to determine the temporal response of those microbes most likely to engage in octanoic acid degradation in AD sludge. This method can be applied to the identification of specialist microbes with a role in degradation of a range of other compounds in AD, enhancing our understanding of microbial community interactions and facilitating the development of strategies for process optimisation.

Table of contents

List of Tables

List of Figures

Acknowledgements

First and foremost, I am extremely grateful to God for the good health and wellbeing that were necessary to complete this thesis.

Declaration

I, Yumechris Amekan, declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

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

a) The library preparation and metagenomic sequencing, as described in Chapter 2.5.3.1 and 2.5.3.2, was performed by Dr. Sally James.

b) The LC-MS/MS and spectral search against proteins predicted from MAGs, as described in Chapter 2.5.4 and 2.7, was performed by Dr. Adam Dowle.

c) The bioinformatic analysis, as described in Chapters 2.3.6, was performed by Annabel Cansdale and Dr. Sarah Forrester.

d) The link between the associated taxonomically classified MAGs with the acquired prokka accessions for taxa identification, as described in Chapter 2.7 was performed by Dr. Sarah Forrester.

1 Introduction

1.1 Anaerobic digestion for renewable energy generation and waste management

Fossil fuels (crude oil, coal and natural gas) supply approximately 80% of global primary energy demand. This grew at an average rate of 2.1% per year between 2009 – 2019 (Figure 1.1; British Petroleum Company, 2020). Fossil fuels have helped to propel the growth and development of the world's economy, technology and quality of life. However, fossil fuel combustion emits around 30 Gt of carbon dioxide (CO_2) and other greenhouse gasses (GHG) per year (IEA, 2019; British Petroleum Company, 2020), making it the major contributor to global warming and climate change. Moreover, political turmoil and military conflict in several fossil fuel producing countries also pose a challenge to global energy security. Therefore, the development and application of alternative energy sources are needed to lessen global dependence on fossil fuels, ensure sustainable energy development and reduce adverse impacts on the environment.

Biofuels are one of the most reliable and sustainable energy alternatives, which support the generation of renewable energy (electricity and gas) from organic materials (biomass) (Dahiya, 2014). Biofuels play an important role in the transition towards low carbon energy systems by reducing reliance on fossil fuels and can contribute to global climate change mitigation. Globally, biofuels made up 1% of world energy supply in 2019 and have increased 9.2% annually from 2009 to 2019 (Figure 1.1). In the United Kingdom (UK), biofuels contribution has grown by 24.2% per annum since 2009 and, when combined with other renewable energy

sources, helped reduce CO_2 emissions by 2.5% annually in 2009 – 2019 (Figure 1.1-4).
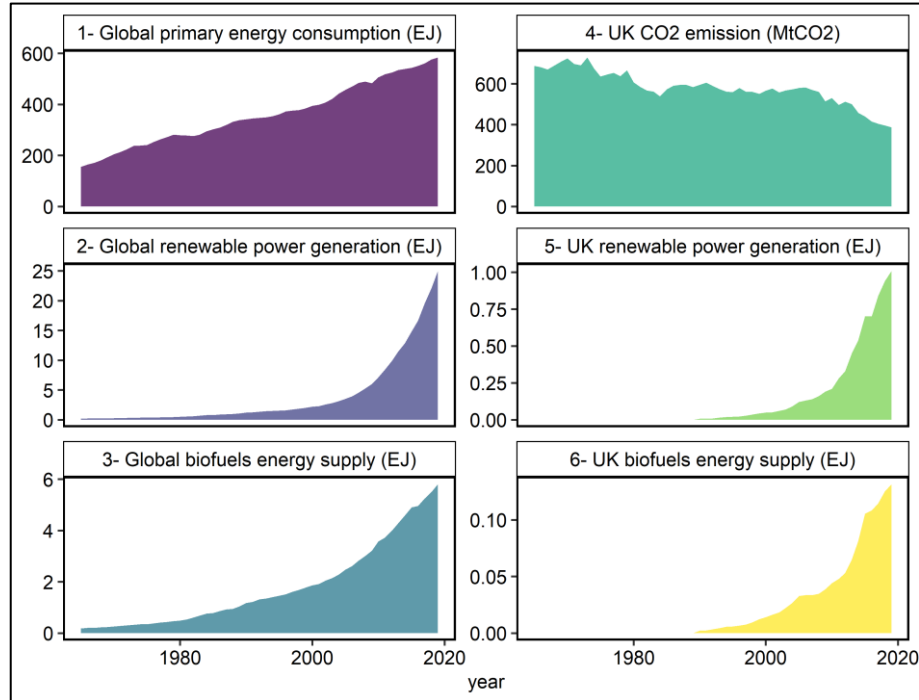
Figure 1.1. Bioenergy growth and contribution to climate change mitigation. (1 – 3) Annual total energy consumption, renewable energy generation and bioenergy supply globally (4 – 6) profile of UK bioenergy generation and carbon emissions from 1965 – 2019. Data are taken from BP Statistical Review of World Energy 2020. EJ = Exajoule, is unit of energy equal to 10^{18} joules; $MtCO_2$ = Metric tons of carbon dioxide equivalent.

   One of the most prevalent bioenergy technologies worldwide is anaerobic digestion (AD), converting organic waste into biogas (Angenent et al., 2004; Werner et al., 2011; Vanwonterghem et al., 2014) and a nitrogen-rich digestate that can be used as bio-fertiliser (Lukehurst et al., 2010; Al Seadi et al., 2013). AD has the potential to offer significant environmental and economic benefits via waste treatment. AD is an ancient technology that has been used as early as 900 BC, where biogas was used to heat bathwater in Assyria, and has been widely applied in the UK to treat and recover energy from sewage sludge for over 125 years (Abbasi et al., 2012). AD technology is flexible (Angelonidi and Smith, 2015; Mauky et al.,

2015) and plants can be built at many different scales and deployed across a range of applications. Current applications include on-farm energy recovery from animal manure and plant waste as well as treatment of municipal food waste and wastewater. Additional benefits from these applications include landfill reduction and fewer CO_2 emissions (Nguyen et al., 2007; Liu et al., 2012). Currently, over 100 million tonnes of UK generated biodegradable material from agricultural, industrial, municipal/commercial and sewage sludge are processed through 661 AD sites for biogas production, in order to generate heat, electricity and biomethane to the grid (BtG) (Figure 1.2).
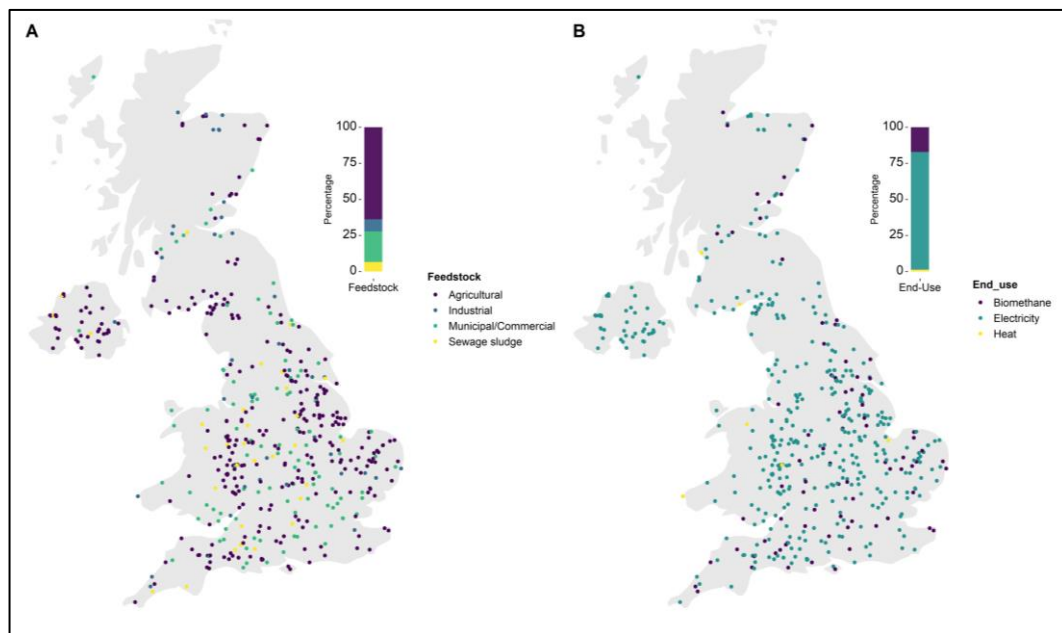
Figure 1.2. Anaerobic digestion plants in the UK. The AD map shows all operational anaerobic digestion plants in the UK based on the type of feedstock (A) and the end-use of the biogas (B). Data are taken from The Anaerobic Digestion and Bioresources Association (ADBA).

1.2 Overview of general AD plant setup and operation

The complete process from receiving the organic waste (feedstock) to biogas utilisation, is facilitated by multiple interconnected units in an AD facility (Figure

1.3). There are numerous details of facility design that must be carefully considered, as per the nature of the feedstock (Table 1.1). The key consideration for achieving optimal biogas yield and efficiency are the solids content, whether the system runs as a batch or continuously, operational temperature, and the number of single/multistage processes (Vasco-Correa et al., 2018).

Table 1.1 Characteristics of common AD feedstock (Zhang et al., 2014; Vasco-Correa et al., 2018).

Feedstock	Characteristics	Biogas yield (m^3/kg VS)	Total solids (%)	C/N ratio
Sewage sludge	- a mud-like residue resulting from wastewater treatment - contains valuable organic matter and nutrients (such as nitrogen and phosphorus), heavy metals and pathogens (viruses and bacteria) - Low digestibility: could be improved by pre-treatment or co-digestion with other feedstocks.	0.8 - 1.2	2 – 30	40 – 70
Food waste	- readily degradable, high moisture content, low pH and high in solubility. - Requires size reduction - High variability in composition. - May generate inhibition through acidification.	0.3 - 0.8	5 – 30	15–30
Agricultural residues and energy crops	- Existing in plentiful supply. - High (ligno)cellulose content - Needs pre-treatment to enhance biodegradability	0.2 - 0.5	20 – 80	40 – 150
Animal manure	- Usually co-digested with bedding material (straw) or another high-carbon biomass - High buffer capacity - Relatively high in ammonia - Rich in nutrients and trace elements	0.1 – 0.6	2 – 20	3 – 15
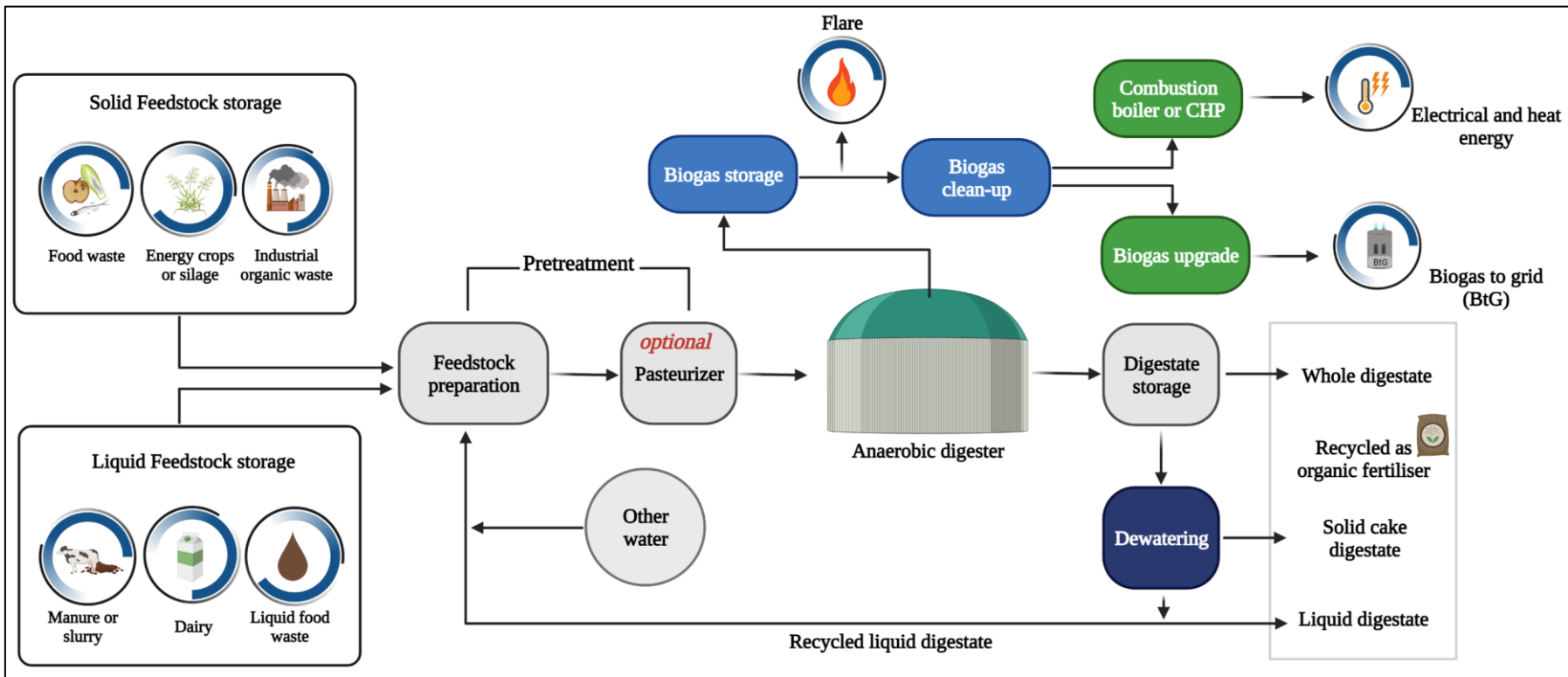
Figure 1.3. Simplified configuration of AD facilities receiving a variety of feedstocks (Pöschl et al., 2010). Created with BioRender.com.

For solids content, there are two broad categories of anaerobic digester operations: as a 'dry' process, in which the feed consistency is dry enough to handle and stack (15 – 30% solids), and as a 'wet' process, where the feed is moist enough to be moved via pipes by a pump (4 – 13% solids) (Rocamora et al., 2020). AD can be performed as a batch or a continuous process. In a batch system, organic material together with biomass from sludge, as inoculum, are placed into a reactor at the beginning and sealed for the duration of the process. Generally, biogas production will follow a normal distribution pattern over time which can be used by the operators to determine when the batch AD process has completed. In continuous systems, organic material is continuously added to the reactor (continuous complete mixed). The digestate (the processed matter after the AD of a biodegradable feedstock) is removed continuously or periodically, resulting in constant production of biogas. A single, or multiple vessels/reactors in sequence, may be used in continuous systems. In single-stage anaerobic systems, all interconnected biochemical processes in AD (see section 1.3) take place simultaneously in a single reactor, while multi-stage systems required generally two reactors to separate and optimise the hydrolysis along with acidogenesis and methanogenesis (Ganesh et al., 2014). Continuous systems require more complex designs, but are probably more cost effective than batch systems. This is because batch digesters need regular and repeated emptying and set-up, have a higher initial construction cost and require larger volume digesters (spread across several batches), to handle the same amount of waste as continuous systems (Abowei et al., 2009; Nicholls, 2015).

Generally, AD can be performed at different temperature ranges, such as psychrophilic (4 – 15 $^{\circ}$C), mesophilic (20 – 40 $^{\circ}$C) and thermophilic (45 – 70 $^{\circ}$C)

(Kim et al. 2017; Náthia-Neves et al. 2018). The optimal operational temperature is usually determined by the feedstock used and its characteristics (Table 1.1; Nichols, 2015). In practice, most digesters are designed to operate in the mesophilic range of 30 – 38 ºC, while some are designed for the thermophilic range of 50 – 57 ºC (Metcalf et al., 1991). This allows optimal growth of the anaerobic bacteria and archaea involved in the breakdown of the organic matter. Thermophilic digestion processes allow for higher loadings with shorter hydraulic retention times (HRTs), higher conversion efficiencies, and pathogen disinfection, whereas mesophilic digestion is more stable, less susceptible to ammonia nitrogen toxicity, and needs less energy input (Lloret et al., 2013; Yu et al., 2014; Moestedt et al., 2014). Some mesophilic AD plants are combined with a thermal hydrolysis process (THP), in which the feedstock is pressure-cooked at high temperature (120 ºC) under pressure to increase feedstock digestibility prior to mesophilic digestion. Whichever thermal regime is used, it has a strong effect on the microbial communities involved in AD, resulting in significant variations in the types and abundance of microbes adapt to, and thrive in, the digester (De Vrieze et al. 2015; Kirkegaard et al. 2017).

AD relies on a diverse microbial community working through a syntrophic series of interrelated biochemical processes (hydrolysis, acidogenesis, acetogenesis, and methanogenesis; see section 1.3). The synergistic interactions within these complex communities influences the speed and efficacy of resource recovery. Although AD has been subjected to substantial process engineering, the underpinning microbial communities have been viewed largely as a 'black box'. Therefore, we need to have a better understanding of how the microbial communities associated with AD interact and function as a single biological machine.

1.3 Interconnected biochemical processes and core microorganisms in anaerobic digestion
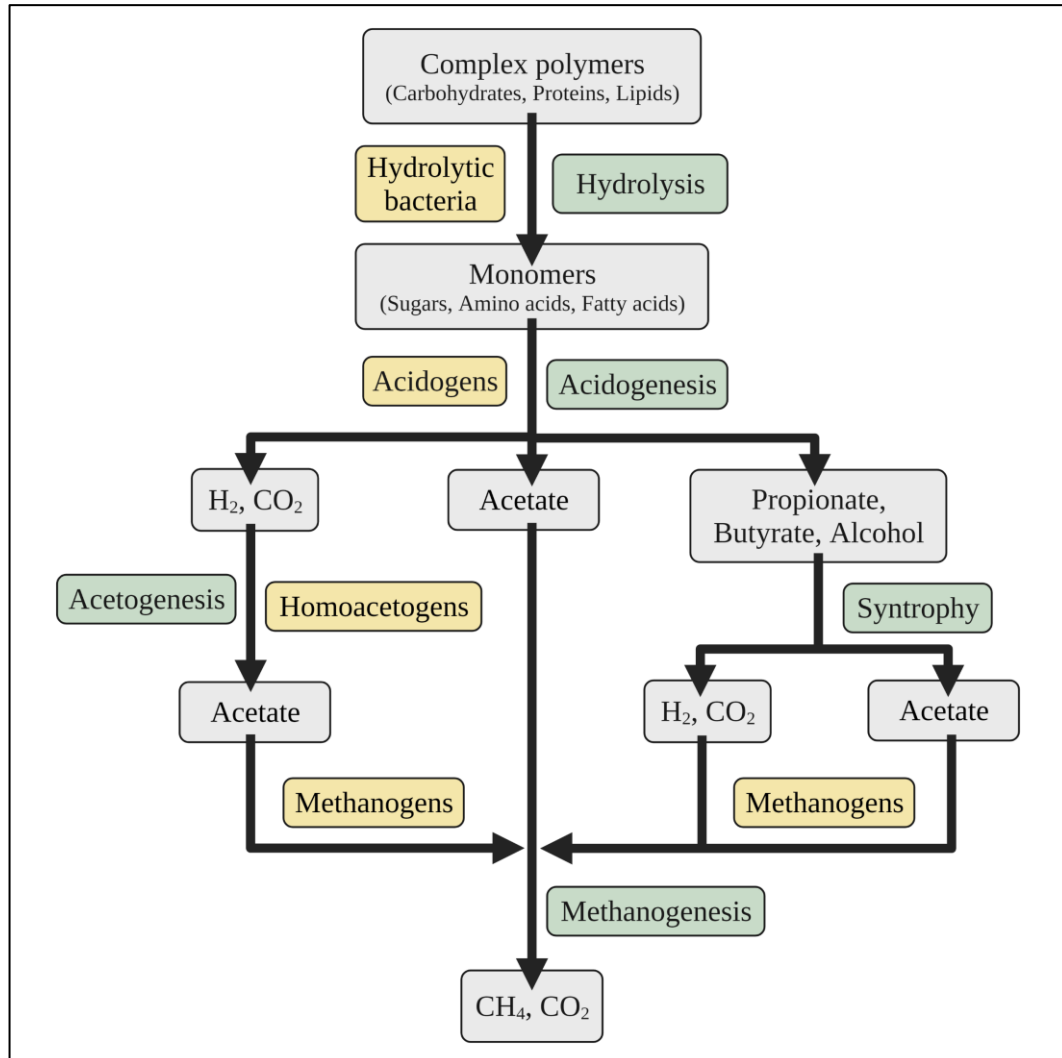
Figure 1.4. The interrelated biochemical functions in AD (Wirth et al., 2012; Heeg et al. 2014; Díaz et al. 2018). Colours denote resource (grey), microbial groups (yellow) and biochemical processes (green). Created with BioRender.com.

Anaerobic degradation of organic waste involves a diverse community of anaerobic bacteria and archaea that sequentially utilise each other's metabolic by-products through a series of interconnected biochemical functions to yield biogas, a gaseous product containing methane (CH_4; $48 - 65\%$), carbon dioxide (CO_2; $15 -$

45%), nitrogen (N_2; $<1 - 17\%$) and other trace gases (Rasi et al., 2007; Bond and Templeton, 2011; Plugge, 2017). Conceptually, the four primary and sequential biochemical processes (Figure 1.4) are (i) hydrolysis of complex polymers into simple soluble products, (ii) acidogenesis (fermentation of simple soluble products into volatile fatty acids (VFAs)), (iii) acetogenesis (anaerobic oxidation of VFAs into acetates and hydrogen), and (iv) methanogenesis (CH_4 production from acetate and hydrogen by methanogenic archaea).

1.3.1   Hydrolysis

Complex organic materials (polymers) that cannot be directly transported across a microorganism's cell membrane, such as polysaccharides, lipids and proteins, are initially hydrolysed into soluble monomers by hydrolases (cellulase, xylanase, pectinase, amylase, lipase, and protease) excreted by hydrolytic bacteria (Angelidaki et al., 2011; De Francisci et al., 2015). Hydrolytic bacteria are very diverse phylogenetically, but many studies support findings that *Firmicutes* and *Bacteroidetes* are the two dominant phyla in AD responsible for the breakdown of polymers (Figure 1.5; De Vrieze et al., 2015; Hassa et al., 2018). *Firmicutes* and *Bacteroidetes* abundance vary depending on AD operational conditions and the type of feedstock present. Commonly found members within the *Firmicutes* and *Bacteroidetes* phyla are the genera, *Clostridium* (O'Sullivan et al., 2005; Zverlov et al., 2010) and *Bacteroides* (Kampmann et al., 2012; Hanreich et al., 2013), respectively.
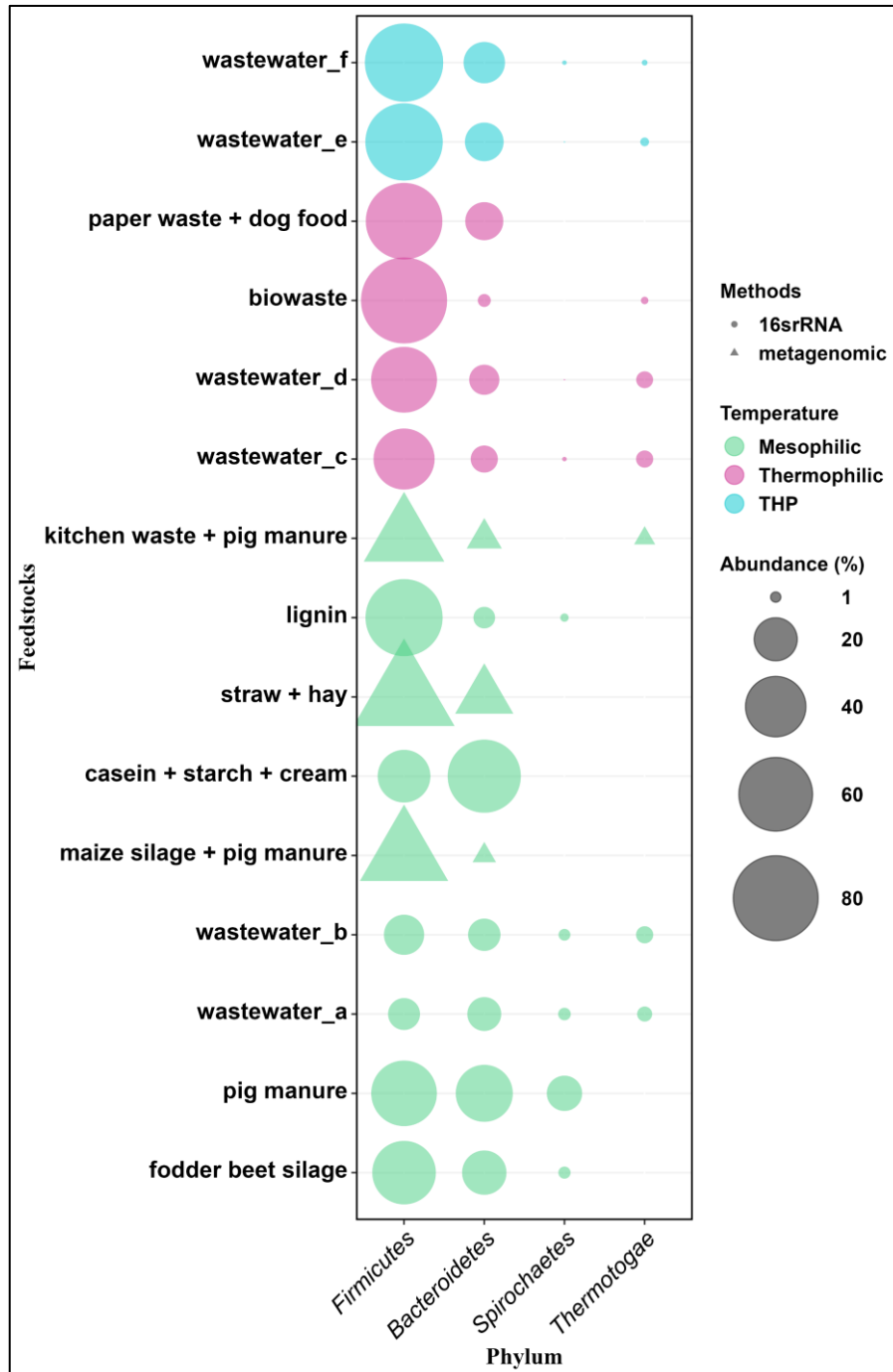
Figure 1.5. Hydrolytic bacterial phyla abundance in anaerobic digesters operated at different temperatures and with different feedstocks. Bubbles are displayed only if the relative abundance is ≥1%. Bubbles are coloured by temperature applied on the AD system, including mesophilic reactors fed with thermal hydrolysis (THP) treated waste. Abundance value were generated based on 16S rRNA (Klocke et al., 2007; Goberna et al., 2009; Liu et al., 2009a; Tang et al., 2011; Kampmann et al., 2012; Wu and He, 2013; Kirkegaard et al., 2017) and metagenomic sequencing (Wirth et al., 2012; Hanreich et al., 2013; Li et al., 2013).

Clostridium is a genus of obligate anaerobes, Gram-positive, rod-shaped and endospore-forming bacteria (Wells and Wilkins, 1996) belonging to *Firmicutes*. Members of this genus are able to secrete a variety of carbohydrate-degrading enzyme, allowing them to directly saccharify a range of polysaccharides (cellulose, starch, xylan, pectin, etc.) and oligosaccharides (cellobiose, lactose, maltose, etc.) in mesophilic and thermophilic AD (Lee et al., 1985). *Clostridia* also exhibit a wide versatility in terms of monosaccharide fermentations, by effectively taking up both hexoses and pentoses simultaneously (Freguia et al., 2013). Thus, in terms of their substrate hydrolysis and uptake capabilities, they are the ideal genus for biofuel production from lignocellulosics. Cellulolytic *Clostridium* (Table 1.2) hydrolyse polysaccharides by direct and specific adhesion to substrate via the cellulosome (Bayer et al., 1998; Desvaux, 2005), an extracellular multi-enzyme complex on the bacterial cell surface containing a scaffolding protein and catalytic subunits displaying a variety of cellulosomal enzymes (Shoham et al., 1999; Blum et al., 2000; Kurokawa et al., 2001; Kataeva et al., 2002). Other *Clostridia* degrade cellulose simply via excretion of a large number of cellulases as single enzymes (Table 1.2; Schwarz, 2001). Furthermore, proteolytic activity is also possessed by a number of Clostridial species, i.e *C. botulinum* (Lee and Riemann, 1970; Sebaihia et al., 2007), *C. proteolyticum* (Jain and Zeikus, 1988), *C. collagenovorans* (Jain and Zeikus, 1988), *C. difficile* (Poilane et al., 1998), *C. thiosulfatireducens* (Hernández-Eugenio et al., 2002), *C. tunisiense* (Thabet et al., 2004) and *C. beihaiense* (Dong et al., 2018), which allows these organisms to hydrolyse proteins into smaller polypeptides or amino acids. Meanwhile, *C. lundense* is exclusively

known as lipolytic *Clostridium* (Cirne et al., 2006), decomposing lipids to directly source carbon by this route.

Table 1.2 Cellulolytic and non-cellulolytic *Clostridium* in anaerobic systems. Optimum AD growth temperature (Temp.; m: mesophilic (20 – 40 ºC); t: thermophilic (45 – 70 ºC), source of isolation and production of cellulosome (Celsm.; + presence; - absence) are indicated.

Species	Temp.	Source	Celsm.	Reference
C. cellobioparus	m	Soil	+	Doi, 2008
C. thermocellum	t	Soil	+	Schwarz, 2001
C. cellulovorans	m	Wood digester	+	Tamaru et al., 2000
C. papyrosolvens	t	Papermill	+	Doi, 2008
C. cellulolyticum	m	Compost	+	Desvaux, 2005
C. josui	t	Compost	+	Sakka et al., 2010
C. acetobutylicum	m	Soil	+	Sabathé et al., 2002
C. aldrichii	m	Wood digester	-	Yang et al., 1990
C. cellulofermentans	m	Manure	-	Yanling et al., 1991
C. herbivorans	m	Pig intestine	-	Varel et al., 1995
C. stercorarium	t	Compost	-	Schwarz et al., 1995
C. hungatei	m	Soil	-	Monserrate et al., 2001
C. clariflavum	t	Anaerobic sludge	-	Shiratori et al., 2009
C. caenicola	t	Anaerobic sludge	-	Shiratori et al., 2009
C. indicum	m	Anaerobic sludge	-	Gundawar et al., 2019

Bacteroides is a genus of obligate anaerobes, Gram-negative and non-endospore-forming bacilli belonging to *Bacteroidetes* (Madigan et al., 2018). Members of the *Bacteroides* also perform polymer hydrolysis in AD alongside *Clostridium*. One such strategy that is used by *Bacteroides* species is the starch utilization system, referred to as the (Sus)-like systems, in which a series of outer-membrane and periplasmic proteins act together to bind, enzymatically degrade and import polyssaccharide products (Martens et al., 2009). To date, few species of polysaccharide hydrolysing *Bacteroides* have been isolated and characterised from

anaerobic digesters, including *Bacteroides xylanolyticus* (Scholten-Koerselman et al., 1986), *B. propionicifaciens* (Ueki et al., 2008), *B. graminisolvens* (Nishiyama et al., 2009), *B. paurosaccharolyticus* (Ueki et al., 2011) and *B. luti* (Hatamoto et al., 2014), whereas the thermophilic *Bacteroides sp.* strain P-1 is known to possess a cellulosome-type multienzyme complex (Ponpium et al., 2000). *Bacteroides* species also exhibit proteolytic activity during AD, i.e. *B. amylophilus* (Blackburn, 1968), *B. ruminicola* (Hazlewood and Edwards, 1981) and *B. fragilis* (Gibson and MacFarlane, 1988). Some *Bacteroides*, e.g. *B. coprosuis* (Whitehead et al., 2005) show esterase/lipase activity for decomposing lipids.

Previous studies have highlighted the importance of polymer structure and the metabolic capability of the microbial communities in the depolymerisation and the hydrolytic activity of polymer substrate fed AD systems (Adney et al., 1991; Parawira et al., 2005; Odnell et al., 2016). It suggests that there are fundamental connections between the microbial communities' composition and their enzymatic functions. The structural specificities of hydrolytic bacteria's extracellular (free and cell-bound) enzymes then determines the complexity and availability of low molecular weight substrates for uptake by the wider AD communities (Arnosti, 2011). Thus, the ability of the hydrolytic microbes to utilise more energy-efficient substrate-uptake mechanisms (Figure 1.6) can influence which phyla or genera proliferate and persist in the digester (Koropatkin et al., 2012).
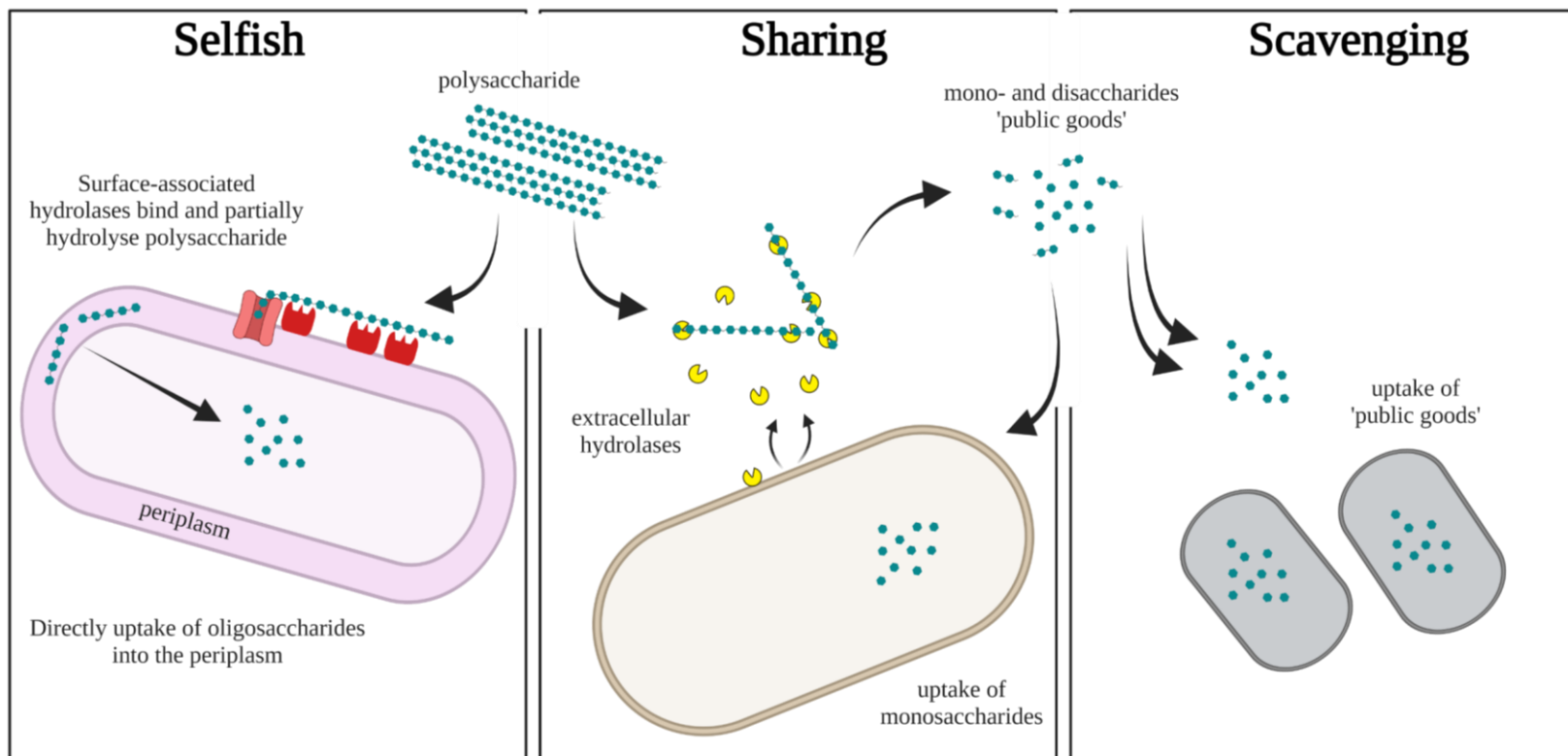
Figure 1.6. Schematic diagram of three main mechanisms of substrate uptake during hydrolysis in AD. Selfish: cells utilise surface-associated enzymes to bind and partially breakdown polymers, which are subsequently carried up into the periplasm for further degradation, with scarcely extracellular hydrolysis products. Sharing: cells use surface-associated or 'free' extracellular hydrolases to degrade polymers into sizes suitable for uptake and yield extracellular hydrolysis products (public goods). Scavengers: cells that unable to produce enzymes for the polymer' hydrolysis and instead take-up the hydrolysis products produced by other organisms (Reintjes et al., 2017; Arnosti et al., 2018; Reintjes et al., 2019). Created with BioRender.com

Clostridium and *Bacteroides* are expected to be dominant among other hydrolytic bacteria in a digester with high hydrolytic activity. Both genera have evolved to compete efficiently for nutrients in highly competitive environments/ecosystems, like AD systems, by utilising 'selfish' uptake mechanisms using substrate processing via surface-associated enzymes (Martens et al., 2009; Reintjes et al., 2019). In *Clostridium*, this is done using the cellulosome and in *Bacteroides* through the use of (Sus)-like systems (Martens et al., 2009). Both systems have been shown to bind and partially degrade polymer substrate without diffusive loss of substrate to the environment, in both gut and marine environments (Cuskin et al., 2015; Reintjes et al., 2019). These 'selfish' mechanisms enable *Clostridium* and *Bacteroides* to be more resource efficient and outcompete other hydrolytic bacteria in the digester that use either 'sharing' (extracellular enzyme producer) or 'scavenging' (non-enzyme producer) approaches to up-take of hydrolysis products as growth substrates (Figure 1.6). The high abundance of these two genera in densely populated AD microbial communities treating polymer substrates (i.e. lignocellulosic material) has been illustrated multiple times (Sun et al., 2015; Soares et al., 2018; Tukanghan et al., 2021).

1.3.2 Acidogenesis and Acetogenesis

Acidogenesis takes hydrolysis products and ferments them. This creates VFAs (e.g. acetate, propionate, butyrate, and valerate), as well as a range of volatile gases (e.g. carbon dioxide, hydrogen and ammonia), whilst generating ATP (Agler et al., 2011; Mathai et al., 2015). Acidogenic bacteria (acidogens) include hydrolytic bacteria as well as fermentative bacteria, which do not produce extracellular

hydrolases and rely upon hydrolytic bacteria for accessible metabolites. *Actinobacteria*, *Bacteroidetes*, *Cloacimonetes*, *Chloroflexi*, *Firmicutes*, and *Proteobacteria* are common bacterial phyla within AD that incorporate species of acidogens (Figure 1.7; Sträuber et al., 2012; Rinke et al., 2013; De Vrieze et al., 2015; Guo et al., 2015; Kirkeegard et al., 2017; Hassa et al., 2018).
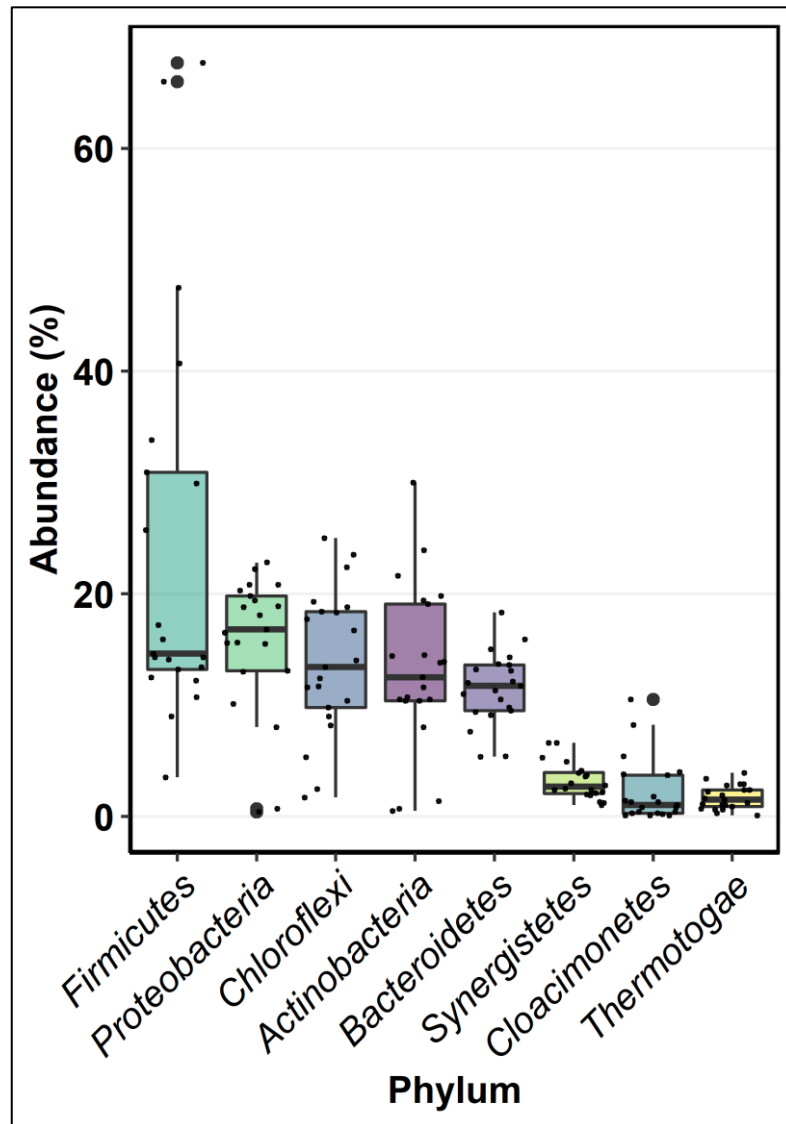
Figure 1.7. Abundances of most prevalent phyla in AD that incorporate species of acidogens. Tukey's box plots show the relative abundance for each phylum, with each box extending from the 25th to the 75th percentile of the distribution. Median indicated by a central line. Data generated based on metagenomics (Guo et al., 2015) and 16s rRNA sequencing (Kirkegaard et al. 2017).

Acidogenesis progresses rapidly upon initiation of hydrolysis (Meegoda et al., 2018). Some end products of acidogenesis, such as acetic acid and H_2/formic acid, can be used directly by methanogens (acetoclastic and hydrogenotrophic) for biogas production, but other intermediates (such as $\geq$ 3-Carbon VFAs and simple alcohols) require further metabolic conversion to the necessary substrates for methanogenesis (Fu et al., 2018; Lemaire et al., 2020). The metabolism of VFAs and other simple alcohols into acetate, formate and CO_2/H_2 in AD relies on syntrophic interaction between acetogens and acetoclastic/hydrogenotrophic methanogens (Mathai et al., 2015; Ziels et al., 2015) via interspecies electron exchanges by means of interspecies hydrogen transfer (IHT) and interspecies formate transfer (IFT) (Boone, 1985; Schink, 1997; Drake et al., 2008; Saha et al., 2020). Cohabitation with autotrophic methanogens allows acetogens to overcome the thermodynamic barrier created by high H_2 partial pressure, since the methanogens metabolise available hydrogen sufficiently rapidly to maintain low hydrogen concentrations overall. Under sufficiently low H_2 concentration conditions acetogenesis becomes thermodynamically favoured (de Bok et al., 2005; Kirchman, 2018). In the case of IFT, four molecules of formate are oxidized to CO_2 by formate dehydrogenase (FDH). One molecule of CO_2 is further reduced to methane (Liu and Whitman, 2008).

Direct interspecies electron transfer (DIET) has been intensively studied as an alternative syntrophic metabolism to IHT and IFT in AD (Dubé and Guiot, 2015). In DIET, electron-donor and electron-acceptor microbes exchange electrons via electrical conduits on the cell membrane, like extracellular c-type cytochromes (e.g. *OmcS*, *MacA*, *OmcC*, *PgcA*) and conductive proteinaceous filamentous

structures/nanowires (i.e. pili), which may allow engagement over long distances (up to 10 mm) (Nielsen et al., 2010; De Vrieze and Verstraete, 2016; Saha et al., 2020; Zhao et al., 2020). Electron exchange via conductive nanowires has been observed as the main mechanism during methane production via co-cultures of *Pelotomaculum thermopropionicum*/*Methanotermobacter thermautotrophicus* (Kouzuma et al., 2015), *Geobacter metallireducens*/*Methanosaeta harundinacea* (Rotaru et al., 2014a) and *Geobacter metallireducens*/*Methanosarcina barkeri* (Rotaru et al., 2014b). Syntrophic acetogens found in anaerobic digesters (Table 1.3) include species in the genus *Smithella*, *Syntrophobacter*, and *Pelotomaculum* for propionate oxidation (3-carbon VFA; Liu et al., 1999; de Bok et al., 2001; Imachi et al., 2007) and *Syntrophus* and *Syntrophomonas* for the oxidation of butyric and longer chain fatty acids ($\geq$ 4-carbon VFAs; Jackson et al., 1999; Sousa et al., 2007a).

Table 1.3 VFA degrading syntrophic microbes in anaerobic system. Scope of trial substrates includes C_3 Propionic, iC_4 *Iso*butyric, C_4 Butyric, iC_5 *Iso*pentanoic, C_5 Pentanoic, iC_6 *Iso*hexanoic, C_6 Hexanoic, C_7 Heptanoic, C_8 Octanoic, C_9 Nonanoic and C_{10} Decanoic acid. The substrates utilisation is indicated by + (utilised), - (not utilised) and ND (not detected or not determined).

Syntrophic acetogens	Substrate utilisation in co-culture with a syntrophic partner											Syntrophic partner	Ref.
	C_3	iC_4	C_4	iC_5	C_5	iC_6	C_6	C_7	C_8	C_9	C_{10}		
Syntrophobacter wolinii	+	-	-	-	-	-	-	-	-	-	-	*Desulfovibrio sp.*	Boone and Bryant, 1980
S. pfennigii	+	-	-	-	-	-	-	-	-	-	-	*Methanospirillum hungatei*	Wallrabenstein et al., 1995
S. fumaroxidans	+	-	-	-	-	-	-	-	-	-	-	*M. hungatei*	Harmsen et al., 1998
S. sulfatireducens	+	-	-	-	-	-	-	-	-	-	-	*M. hungatei*	Chen et al., 2005
Smithella propionica	+	-	-	-	-	-	-	-	-	-	-	*M. hungatei*	Liu et al., 1999
Pelotomaculum thermopropionicum	+	-	-	-	-	-	-	-	-	-	-	*Methanothermobacter thermautotrophicus*	Imachi et al., 2002
P. schinkii	+	ND	-	ND	-	ND	-	-	-	-	-	*M. hungatei*	de Bok et al., 2005
Syntrophomonas wolfei	-	-	+	-	+	-	+	+	+	ND	ND	*M. hungatei; Desulfovibrio sp. G11; Methanobacterium formicicum*	McInerney et al., 1981; Wu et al., 2007; Narihiro et al., 2016
S. sapovorans	-	-	+	-	+	-	+	+	+	+	+	*M. hungatei*	Roy et al., 1986
S. bryantii	-	-	+	-	+	-	+	+	+	+	+	*M. hungatei; Desulfovibrio sp.*	Stieb and Schink, 1985; Wu et al., 2006a
S. cellicola	-	-	+	-	+	-	+	+	+	+	ND	*M. formicicum; Desulfovibrio strain G11*	Wu et al., 2006a
S. curvata	-		+		+		+	+	+	ND	+	*M. formicicum*	Zhang et al., 2004

Species	1	2	3	4	5	6	7	8	9	10	11	Partner	Reference
S. erecta	-	-	+	-	+	-	+	+	+	ND	+	*M. hungatei; M. formicicum*	Zhang et al., 2005; Wu et al., 2006b
S. palmitatica	-	-	+	-	+	-	+	+	+	ND	ND	*M. hungatei*	Hatamoto et al., 2007
S. saponavida	-	-	+	-	+	-	+	+	+	+	+	*Desulfovibrio sp.*	Lorowitz et al., 1989; Wu et al., 2007
S. zehnderi	-	-	+	-	+	-	+	+	+	+	+	*M. formicicum*	Sousa et al., 2007a; Cavaleiro et al., 2010
Syntrophus aciditrophicus	-	-	+	-	ND	-	+	ND	+	ND	ND	*M. hungatei; Desulfovibrio sp. G11*	Jackson et al., 1999
Syntrophothermus lipocalidus	-	+	+	-	+	-	+	+	+	+	+	*M. thermoautotrophicum*	Sekiguchi et al., 2000
Thermosyntropha lipolytica	-	-	+	-	+	-	+	+	+	+	+	*Methanobacterium strain JW/VS-M29*	Svetlitshnyi et al., 1996
Algorimarina butyrica	-	+	+	ND	-	ND	-	-	-	-	-	*Methanogenium AK-3*	Kendall et al., 2006
Strain GraIva1	ND	ND	ND	+	ND	ND	ND	ND	ND	ND	ND	*Desulfovibrio sp.*	Stieb and Schink, 1986

In principle, long-chain fatty acids (LCFAs; ≥13-Carbon), medium-chain fatty acids (MCFAs; 7 – 12-Carbon) and short-chain fatty acids (SCFAs; 3 – 6-Carbon; Schönfeld and Wojtczak, 2016) are oxidized via the β-oxidation pathway (Figure 1.8). Even-chain fatty acids (ECFAs) are broken down via 2-Carbon VFA units (or acetic acid) resulting in the release of H_2 whilst odd-chain fatty acids (OCFAs) eventually produce propionate along with acetate and H_2 (Sousa et al., 2007b; Sousa et al., 2009). Oxidation metabolism is initiated by the activation of fatty acids to fatty acyl-CoA molecules by the enzyme, acyl-CoA synthetase. Fatty acyl-CoA molecules then degrade through β-oxidation (Figure 1.8). This consists of dehydrogenation (mediated by acyl-CoA dehydrogenase), hydration (catalysed by enoyl-CoA hydratase), oxidation (catalysed by 3-hydroxyacyl-CoA dehydrogenase) and thiolytic cleavage (mediated by beta-ketothiolase), which cleaves the terminal acetyl-CoA group and forms a new, shorter acyl-CoA (two carbons shorter than the original fatty acyl-CoA). The shortened acyl-CoA then re-enters the β-oxidation pathway (DiRusso et al., 1999; Sousa et al., 2009). Some iso-forms, e.g. *iso*-butyric (*i*C4) and *iso*-pentanoic (*i*C5), are produced as intermediates during the anaerobic degradation of fatty acids. Anaerobic degradation of fatty acid isomers (branched chain) might firstly be isomerised into their respective normal form (straight chain) before going through β-oxidation (Stieb and Schink, 1989; Wu et al., 1994; Sekiguchi et al., 2000). The isomerisation of butyric and *iso*-butyric acid was catalysed by a butyryl-CoA:isobutyryl-CoA mutase which depended strictly on the presence of coenzyme B12 (Matthies and Schink, 1992).
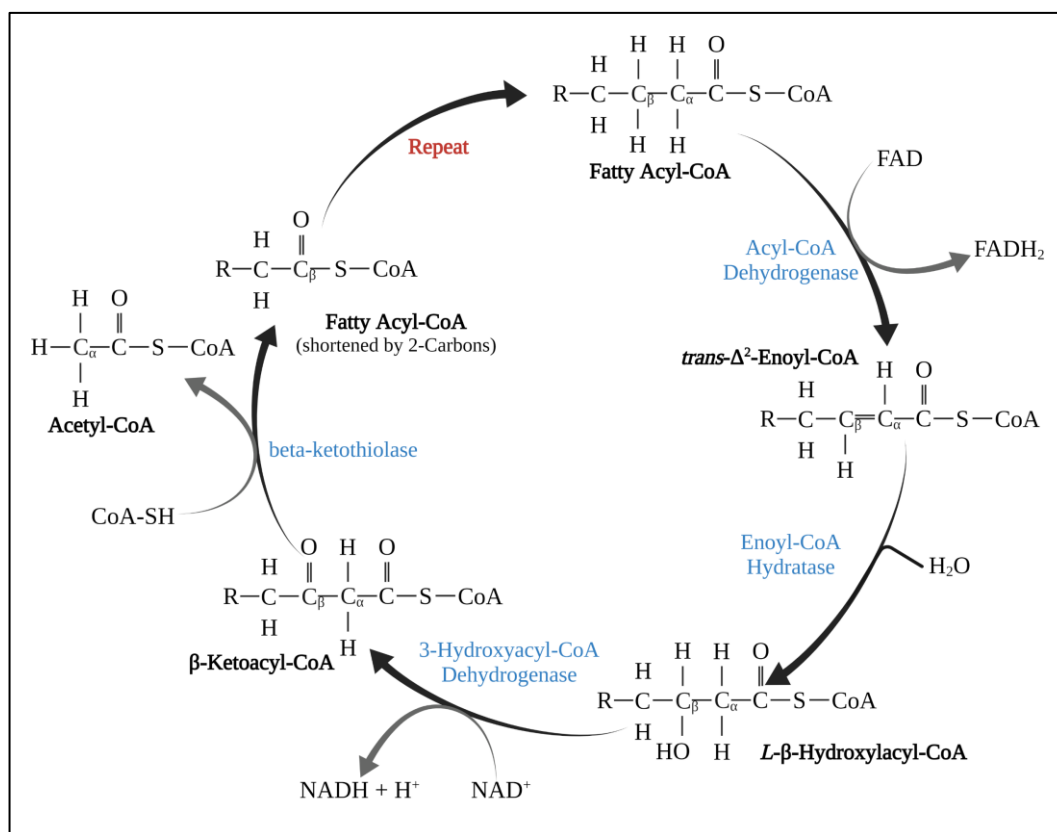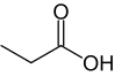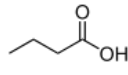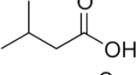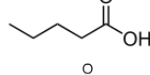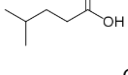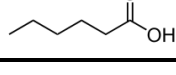
Figure 1.8 β-oxidation of fatty acids. Colours denote compounds (black) and enzymes (blue). Adapted from Schulz et al. (1991). Created with BioRender.com.

SCFAs (< 6-Carbon VFAs; Table 1.4) are important intermediates and metabolites in AD that have been well established as an essential indicator for monitoring the AD process (Schönfeld and Wojtczak, 2016). In comparison to other typical indicators such as pH, alkalinity, gas production and gas composition, they provide rapid and reliable information on AD performance (Hill and Holmberg, 1988; Ahring et al., 1995; Siedlecka et al., 2008). Many analytical methods have been developed to measure and monitor SCFAs present in AD samples, such as distillation, colorimetry, chromatography and various titration techniques (Lahav and Morgan, 2004; Siedlecka et al., 2008). In the distillation method, the organic acids are isolated from the matrix by direct distillation or steam-distillation (a multistage continuous distillation process where steam is used as a stripping gas to

extract the VFAs) which are then titrated with standard alkaline solution. Accurate measurement of VFAs in AD samples using the distillation technique is primarily dependent on matrix recovery (Siedlecka et al., 2008). Titration methods are widely acknowledged to be better for on-site regular monitoring and control in terms of simplicity, speed and cost-effectiveness to measure the total VFA concentrations (Feitkenhauer et al., 2002; Madsen et al., 2011). However, titration methods cannot distinguish the individual SCFAs in the sample because of their similarity in pKa values (Table 1.4). Colorimetric determination of VFAs, is based on the esterification of VFAs present in the sample, followed by the measurement of the esters by the ferric hydroxamate reaction which is known as the Montgomery method (Montgomery et al., 1962). This method is relatively simple, requires commonly available reagents and results in a short analysis time. It is also performed directly on the sample and needs only small sample volumes (0.5 mL). However, the colorimetric method cannot be used to determine individual VFA concentrations. Gas chromatography with flame-ionisation detection (GC-FID), which is mostly used in industrial and research facilities, allows the separation and quantification of individual VFAs based on boiling points difference (Table 1.4; Brondz, 2002; Siedlecka et al., 2008; Madsen et al., 2011). During GC separation, the sample is vaporised and carried across the column by the mobile gas phase (i.e., the carrier gas). VFAs are separated based on their relative vapour pressures and affinities for the stationary phase. The measurements are quantified using a linear fit to a standard curve that is generated from either a manually mixed or commercially available standard solution (Khotsena and Potivichayanon, 2020). This technique requires a great deal of expertise and extensive sample preparation

(centrifugation, filtration and acidification) before injection of the purified sample as AD sample are complex and often viscous or contain high-solids (Ghidotti et al., 2018).

Table 1.4 Short-chain fatty acids (SCFAs) properties and structures (Siedlecka et al., 2008).

SCFAs	Formula	Molar mass (g.mol^{-1})	Boiling point (^{o}C)	pKa	Structure
Formic acid	CH_2O_2	46.025	100.5	3.745	
Acetic acid	$C_2H_4O_2$	60.052	118 – 119	4.756	
Propionic acid	$C_3H_6O_2$	74.079	141.15	4.88	
Isobutyric acid	$C_4H_8O_2$	88.11	155	4.86	
Butyric acid	$C_4H_8O_2$	88.106	163.75	4.82	
Isopentanoic acid	$C_5H_{10}O_2$	102.13	176.5	4.80	
Pentanoic acid	$C_5H_{10}O_2$	102.133	185	4.82	
Isohexanoic acid	$C_6H_{12}O_2$	116.16	200.5	4.84	
Hexanoic acid	$C_6H_{12}O_2$	116.160	205.8	4.88	

1.3.3 Methanogenesis

Methanogenesis marks the final stage of AD by converting end products of acidogenesis and acetogenesis into methane (CH_4; Ferry, 2010). This obligate anaerobic metabolic stage is carried out by methanogenic archaea, a group of organisms commonly known as methanogens (phylogenetically belonging to the phylum Euryarchaeota; Whitman et al., 2001). In AD communities, methanogens have a relatively low abundance (<5% of total reads by 16S rRNA gene approach;

Sundberg et al., 2013; Kirkeegard et al., 2017) but have display high activity, representing up to 30% of total transcript or proteins, based on metatranscriptomics or metaproteomics approaches (Zakrzewski et al., 2012; Hanreinch et al., 2012; Hanreich et al., 2013).

Methanogens can be split into hydrogenotrophic, methylotrophic and acetoclastic groups, based on the main substrate used to generate methane (Table 1.5; Garcia et al., 2000; Thauer et al., 2008). Hydrogenotrophic methanogens are represented by six orders (i.e. *Methanomicrobiales*, *Methanopyrales*, *Methanocellales*, *Methanococcales*, *Methanosarcinales* and *Methanobacteriales*) and almost all species depend on CO_2 reduction to CH_4, with H_2 as the electron donor, but some species oxidise other electron sources (e.g., formate; $HCOO^-$) to form methane (Garcia et al., 2000; Enzmann et al., 2018; Tao et al., 2019). Methylotrophic methanogens are capable of producing CH_4 from various methylated compounds (e.g. methanol (CH_3OH), methylamines ($(CH_3)NH_2$) or dimethylsulfide ($(CH_3)_2S$)). Examples of methylotrophic methanogens are found in the orders *Methanomassiliicoccales*, *Methanobacteriales* and *Methanosarcinales* (Iino et al., 2013; Enzmann et al., 2018). Acetoclastic methanogens, specialist methylotrophic methanogens, utilise the methyl group of acetate (CH_3COO^-) to form CH_4 and generate energy for growth. Acetoclastic methanogens all belong to the *Methanosarcinales*, specifically species of the genera *Methanosarcina* and *Methanothrix* (formerly known as *Methanosaeta*) (Oren, 2014; Welte and Deppenmeier, 2014).

Table 1.5 Reactions for methanogenesis (Garcia et al., 2000; Thauer et al., 2008).

Substrate	Reaction
Hydrogen	$4H_2 + CO_2 \rightarrow CH_4 + 2H_2O$
Formic acid	$4HCOOH \rightarrow CH_4 + 3CO_2 + 2H_2O$
Acetic acid	$CH_3COOH \rightarrow CH_4 + CO_2$
Methanol	$4CH_3OH \rightarrow 3CH_4 + CO_2 + 2H_2O$
Monomethylamine	$4(CH_3)NH_2 + 2H_2O \rightarrow 3CH_4 + CO_2 + 4NH_3$
Dimethylamine	$2(CH_3)_2NH + 2H_2O \rightarrow 3CH_4 + CO_2 + 2NH_3$
Trimethylamine	$4(CH_3)_3N + 6H_2O \rightarrow 9CH_4 + 3CO_2 + 4NH_3$
Dimethylsulfide	$2(CH_3)_2S + 2H_2O \rightarrow 3CH_4 + CO_2 + H_2S$

Methanogen diversity and activity varies across AD systems with substrate and temperature (Abendroth et al., 2015; Kirkeegard et al., 2017). Hydrogenotrophic methanogens, like *Methanoculleus*, appear to be dominant in the digesters treating mixed organic wastes or co-digesters (Abendroth et al., 2015; Kern et al., 2016) and CO_2 biomethanisation digesters where hydrogen is injected into the digester to react with CO_2 in the biogas to create more CH_4 (Tao et al., 2019), whereas acetoclastic methanogen *Methanothrix* (*Methanosaeta*) is prevalent in digesters treating sewage sludge (Abendroth et al., 2015). *Methanosarcina*, a genus which possesses all three known pathways for methanogenesis, is plentiful in AD systems treating leachate from leach-beds (Kern et al., 2016). The two acetoclastic methanogen genera, *Methanosarcina* and *Methanothrix*, utilise identical substrate but coexist and occupy different niches in AD due to differences in enzyme affinity and growth kinetics. *Methanosarcina* thrives under high levels of acetate (>1 mM) due to its high maximum rate of acetate utilisation (k; 12.2 mg COD/mg VSS.d), specific growth rate (μ_{max}; 0.3 d^{-1}) and high half-saturation coefficient (K_S; 280 mg COD/L), while *Methanothrix* has greater substrate affinity and therefore is

prevalent in digesters with low acetate (<1 mM) availability because of its low k (10.1 mg COD/mg VSS.d), μ_{max} (0.1 d^{-1}) and low K_S (49 mg COD/L) (Gujer and Zehnder, 1983; Conklin et al., 2006; Karakashev et al., 2005; Razaviarani and Buchanan, 2014).

Mesophilic (~37 ^{o}C) and thermophilic (~55 ^{o}C) anaerobic digesters develop distinct methanogen communities. Mesophilic digesters are dominated by *Methanothrix* but a variety of hydrogenotrophic methanogens (i.e *Methanolinea*, *Methanospirillum*, and *Methanobrevibacter*) are commonly found at lower population densities. Thermophilic digesters are dominated by the hydrogenotrophic methanogen *Methanothermobacter*, with less abundant populations of *Methanosarcina*, *Methanobrevibacter*, and *Methanothrix* (Kirkeegard et al., 2017).

1.3.4   Rate-limiting steps in the AD process

AD metabolism of complex substrates is a multi-step process (section 1.3.1 – 1.3.3) that is kinetically controlled by the slowest step, termed the rate-limiting step (Hill and Root, 1977). In AD, rate-limiting steps are determined by substrate characteristics (particulate or soluble), temperature, loading rate and microbial community ratio (Speece, 1983; Pavlostathis and Giraldo-Gomez, 1991; Ma et al., 2013).

Hydrolysis is often found to be the rate-limiting step in the AD of particulate biopolymers (carbohydrates, proteins and lipids). Hydrolytic activity is commonly assumed to follow first-order kinetics with respect to the concentration of degradable particulate organic matter. The rate of hydrolysis determines the

potential maximum substrate concentration for methanogens for a given retention time, which in turn determines maximum possible methanogen specific growth rate (Pavlostathis and Giraldo-Gomez, 1991). For each complex substrate (Table 1.6), apparent hydrolysis rate constants (k_h, d^{-1}) in mesophilic digesters (> 12 days retention time) varies as follows: lipids, 0.04 – 1.7 d^{-1} proteins, 0.01 – 0.03 d^{-1} cellulose, 0.04 – 0.62 d^{-1} and hemicellulose, 0.54 d^{-1} (O'Rourke, 1968, as cited in Pavlostathis and Giraldo-Gomez, 1991; Gujer and Zehnder, 1983). In general, protein hydrolysis under anaerobic conditions is slowest compared to other biopolymers.

Temperature has a strong effect on the hydrolytic rate of complex biopolymers (Table 1.6). In psychrophilic digesters (<20 ºC, >12 days retention time), anaerobic degradation of lipids becomes nil, with no change in protein hydrolysis rate, while cellulose hydrolysis rate is 3 – 6 times slower than in mesophilic digesters (O'Rourke, 1968, as cited in Pavlostathis and Giraldo-Gomez, 1991). Moreover, some reports show that the inoculum-substrate ratio (ISR on a volatile solids (VS) basis) significantly affects biogas production rate in AD utilising complex substrates. ISR ≥2 increases hydrolysis and methanogenesis, but ISR of 1 is recommended to avoid inhibition of AD functional processes (Chen and Hashimoto, 1996; Lopes et al., 2004; Liu et al., 2009b; Raposo et al., 2009).

Table 1.6 Hydrolysis rate constant (K_h, d^{-1}) of polymers in steady state and continuous flow laboratory scale under different temperature (O'Rourke, 1968, as cited in Gujer and Zehnder, 1983).

Polymers	Temperature (^{o}C)	Retention time (days)				
		5	**10**	**15**	**30**	**60**
Lipids	15	-	0	0	0	0
	20	0	0	0.02	0.05	0.03
	25	0	0.01	0.09	0.07	0.03
	35	0.01	0.17	0.11	0.06	0.04
Cellulose	15	-	0.05	0.03	0.10	0.08
	20	0.0.9	0.14	0.13	0.14	0.10
	25	0.29	0.27	0.27	0.34	0.16
	35	1.95	1.21	0.62	0.38	0.21
Protein	15	-	0.03	0.02	0.01	0.01
	20	0.08	0.04	0.03	0.02	0.01
	25	0.09	0.04	0.03	0.02	0.01
	35	0.10	0.05	0.03	0.02	0.01

Methanogenesis is often considered the rate-limiting step in easily degradable (soluble) organic matter-based AD systems (Mosey and Fernandes, 1989; Tomei et al., 2009). In a steady state anaerobic mesophilic digester treating soluble organics, acidogenesis rate (maximum growth rate at 35 ^{o}C, μ_{max} 2.0 d^{-1}; yield, Y_{max} 0.15 Kg VSS/Kg COD; substrate affinity, K_s 0.2 Kg COD/m^3) is faster than methanogenesis rate (μ_{max} 0.4 d^{-1}, Y_{max} 0.03 Kg VSS/Kg COD, K_s 0.05 Kg COD/m^3) (Henze and Harremoës, 1983). This suggests that the rate of CH_4 production in this digester is proportional to the kinetics of methanogenesis. Organic loading rate (OLR (Kg/L/d) on VS or COD basis) appears to determine the rate-limiting step in easily degradable AD systems. At low OLR, the rate-limiting step is expected to be acidogenesis, as evidenced by low VFA concentrations. However, as OLR increases, methanogenesis may gradually become the rate-limiting step, as evidenced by VFA accumulation (Henze and Harremoes, 1983).

1.4 Understanding microbial communities in AD

Over centuries of AD application, there has been considerable process and engineering development, including reactor configuration, categorisation of feedstocks, quantification of methane production per unit of feedstock, mass and energy budget analyses of digestion systems, up-scaling of laboratory systems to commercial reality, and computational modelling (Murphy and Thamsiriroj, 2013; Zhao et al., 2017). However, the underpinning microbial communities (anaerobic bacteria and archaea) that mediate essential AD biochemical processes have traditionally been treated largely as a 'black box', meaning the microbial community structure was only retrospectively investigated in case of failure (Koch et al., 214; Carballa et al., 2015). Recent estimations suggest that a typical millilitre of municipal and industrial wastewater contains $\sim10^{20}$ bacterial and archaeal cells (Flemming and Wuertz, 2019), which suggests that AD reactors are commonly fed with substrates that include a greater range of taxonomic community members and metabolic functions than described in section 1.3. Therefore, it is important to understand the full complexity of the AD microbiome and its available competitive and syntrophic metabolisms, to optimise energy recovery from AD.

The earliest study of the AD microbiome was attempted at the beginning of the 20th century. This study relied on culture-dependent techniques and showed that there were multiple unculturable microbes in AD that could not be isolated and grown in pure cultures (Barker, 1936; McBee, 1954). Over the last three decades, several culture-independent approaches have been developed (see sections 1.4.2 – 1.4.3), assisted by rapid developments in sequencing technology next-generation sequencing, NGS. NGS can provide more comprehensive information at a lower

cost, overcoming limitations of culture-dependent approaches for determining the phylogenetic and functional diversity of AD microbial communities. Currently, 4194 bacterial and 51 archaeal species have been discovered from wastewater treatment and bioenergy systems using culture-dependent and independent approaches (Nierychlo et al., 2020). Nevertheless, the majority of AD microbial community species (~70%) have not yet been characterised (Narihiro et al., 2015) and their potential contribution to AD function requires further investigation.

1.4.1 The importance of considering microorganism status in AD microbiome studies

The AD microbiome encompasses a diversity of taxa and metabolic functions that underpin AD efficiency and stability. A critical, and commonly unanswered, question is whether all the microbes found in the system are active and/or alive. It is possible that observed anaerobic degradation of organic waste is mediated by a small number of living anaerobic bacteria and archaea, while the majority of bacterial/archaeal cells are dead or quiescent/dormant. Quantifying the actual number of live versus dead cells in complex AD communities is difficult. A particular challenge is that microbial cells can be in different metabolic states, ranging from inactive (which includes cells that are both truly dead and those that are merely inactive, awaiting better environmental conditions to re-activate) to active (growing and reproducing) cells. Thus, highly diverse anaerobic microbes in AD may not be equally active at certain times or under different environmental conditions.

Some microbes in AD persist in metabolically inactive states, allowing them to survive unfavourable conditions, such as nutrient starvation or limitation, in toxic chemical concentrations, and under temperature or pressure changes (Dworkin and Shah, 2010). Three strategies are used by microbes to resist growth-limiting stress (Rittershaus et al., 2013). "Bust and boom" refers to the strategy based on dynamic persistence of a small subpopulation (i.e. *Escherichia coli*) that subsists on limited nutrients derived from dead cells and that has the ability to replicate rapidly once growth conditions are conducive. "Cellular quiescence", also known as "viable but non-culturable (VBNC)," refers to metabolically active nonreplicating cells (e.g. *Mycobacterium tuberculosis*, *Vibrio cholerae*, *Salmonella enterica*, and *Legionella pneumophila*) that stop replicating, retain their membrane potential, and do not undergo apparent morphological differentiation when exposed to growth-limiting stress. The ''sporulation'' strategy refers to the production of metabolically inactive spores (i.e. by *Bacillus spp.* and *Clostridium spp.*) upon exposure to adverse growing condition (Dworkin and Shah, 2010; Rittershaus et al., 2013).

Currently, most AD functional processes are ascribed to organisms where taxonomic community composition shows an increased/decreased abundance concurrent with increased/decreased functionality. While informative, abundance is complicated by the coexistence of many different taxa with similar functional characteristics (functional redundancy; Louca et al., 2018) or closely related taxa with very different physiologies and environmental tolerances (Fierer et al., 2012) in AD samples. Microbes in inactive/quiescent states and necromass from recently AD community members can make it difficult to deduce which microbes are actively contribute to AD functions. Some studies provide evidence that changes in

AD microbial abundance/composition does not influence biogas production rates (Fernández et al. 2005; Langer et al. 2015). This suggests that, in some cases, microbial composition is not actively controlling AD function, since microbial systems carry out the interconnected biochemical processes at similar rates regardless of composition differences.

Direct correlation of function and abundance could be problematic in AD complex microbial community analyses. Quantifying abundance is challenging since genetic information is most commonly retrieved via DNA extraction, PCR amplification, sequencing and taxonomic classification (Brooks, 2016). Each step in this process has the potential to introduce bias and significantly alter our perception of the microbiome's true composition. Particularly, in the case of low abundance microbes which are often critical to AD function and stability, it is difficult to accurately determine abundance if sequencing depth is low (Mainali et al., 2017). AD microbiome studies tend to use relative abundance, rather than absolute, which cannot tell you whether a taxon is more or less abundant (the direction of the change) or by how much (the magnitude of the change) between two experimental conditions or samples (Barlow et al., 2020). Furthermore, to date, there is no standardised bioinformatics pipeline to process AD-derived NGS data, so implementation is left to individual researchers, raising concerns about reproducibility (Leach et al., 2012). Inaccuracies in observed microbial proportions/abundances can lead to incorrect conclusions about their effect on AD function.

Our understanding of microbial communities associated with AD currently relies strongly on marker gene profiling and metagenomic sequencing (discussed in

sections 1.4.2 – 1.4.3), which reveals phylogenetic diversity but does not provide information concerning microbial activity or the close associations that form between syntrophic organisms. Moreover, these popular DNA-based sequencing analysis methods cannot exclude or identify extracellular DNA, which affects interpretation of the data. For example, members of the *Chloroflexi* phylum are often found in high abundance in AD systems based on genomic DNA analysis, due to the persistence of their extracellular DNA (Petriglieri et al., 2018). However, they are not metabolically active cells and probably play a minor role in AD (Petriglieri et al., 2018; Speirs et al., 2019). Rather, to advance our understanding of the microbes that power AD sequential biochemical processes, the study of the AD microbiomes requires characterisation and quantification of which microorganisms are active (growing and reproducing), inactive (truly dead), or dormant (resting) (Singer et al., 2017; Emerson et al., 2017).

1.4.2 Marker gene-based (primer-based) approaches for profiling microbial community structure in AD

One culture-independent method that provides insight into the taxonomic diversity of microbes present in AD involves using sequences of the methyl coenzyme-M reductase (*mcrA*; Luton et al., 2002; Wilkins et al., 2015) and the 16S rRNA gene (Woese and Fox, 1977). The 16S rRNA gene is a ribosomal RNA (rRNA) gene (~1500 bp) that encodes the small subunit of ribosomes (SSU rRNA). The 16S rRNA gene contains highly conserved regions (that remain relatively unchanged among different organisms) present in all prokaryotes as well as hypervariable regions (V1-V9) that differ greatly between related microbes allowing specific taxonomic identification (Janda and Abbott, 2007; Kirchman,

2018). The whole 16S rRNA gene, specific regions, and/or the *mcrA* gene can be used as taxonomic marker genes (Kim et al., 2020).

The development of high-throughput sequencing has made it feasible to rapidly sequence thousands of amplicons (PCR-amplified marker gene sequences) in parallel, yielding large datasets at low cost and allowing the inclusion and analysis of low abundance populations (Hassa et al., 2018; Lim et al., 2020). Results are clustered based on similarity, generating operational taxonomic units (OTUs). Representative OTU sequences are compared with reference databases to infer likely taxonomy (Schloss and Handelsman, 2005; Johnson et al., 2019), with thresholds for genus and species level identification being >95% and >97%, respectively (Yarza et al., 2014). Recently, new methods, that assess amplicon sequence variants (ASVs), have been developed to replace OTUs in marker gene data analysis (Eren et al., 2013; Eren et al., 2015; Callahan et al., 2017). ASV methods infer biological sequences in the sample prior to the introduction of amplification and sequencing errors, and distinguish sequence variants differing by as little as one nucleotide (Callahan et al., 2017). ASV methods have demonstrated sensitivity and specificity comparable to or better than OTU methods and better discriminate ecological patterns (Eren et al., 2013; Eren et al., 2015).

Currently, a common approach for studying microbial communities in anaerobic digesters uses a combination of molecular fingerprinting methods and high throughput sequencing of taxonomic marker gene PCR amplicons, for instance DGGE together with Illumina MiSeq sequencing (Walter et al., 2019) or T-RFLP with 454-pyrosequencing (Lv et al., 2019). This combination allows greater resolution of AD microbial communities, with the detection of thousands of OTUs

(Walter et al., 2019), compared to AD microbiome profiling via classical automated Sanger sequencing of 16S rRNA clone libraries, which discovered 69 OTUs (Riviere et al., 2009).

Greater insight into AD microbiome diversity assists researchers in understanding community composition and provides the capability to understand the influence of operational conditions (i.e. feedstock and temperature) on microbial community structure, dynamics, performance, efficiency, and stability. However, PCR-based methods are prone to PCR-amplification bias and potentially leading to misrepresentation of data. To address this problem, PCR-free methods have been proposed (Dowle et al., 2016), but come with significant increase in workload and processing cost (for enrichment, library preparation, and required sequencing coverage; Krehenwinkel et al., 2017). Yet, marker gene-based approaches for AD microbiome profiling are unable to unequivocally differentiate metabolically active cells from those that are dead or quiescent (Lim et al., 2020). This substantive concern limits the utility of NGS method applications in AD microbiome studies.

1.4.3 Metagenomics-based approaches to link function and structure of microbial communities in AD

Although marker gene-based approaches help us to understand AD taxonomic diversity, we cannot fully rely on them to predict functional attributes or the functional diversity of AD communities and their effect on AD performance (Green et al., 2008). It has been proposed that there are a set of organisms in complex AD microbial communities that are capable of utilising a spectrum of resource

acquisition strategies based on their genetic content (Sriswasdi et al., 2017; Louca et al., 2018). For example, besides *Methanosarcina* and some syntrophic acetogens as shown in Table 1.3, *Clostridium cellulovorans* is capable of using carbon sources other than cellulose, including lactose, glucose, galactose, maltose, sucrose, cellobiose, pectin, and xylan (Xin et al., 2019). A technique that exceeds the limitations inherent in taxonomic profiling based on marker gene characterisation is required to obtain targeted information on functional capabilities of AD microbiome species.

Shotgun metagenomic sequencing, which allows direct sequencing of all genetic materials recovered from AD samples, enables the researcher to gain insight not only into microbial diversity and abundance but also into the metabolic potential of microbial communities residing within anaerobic digesters (Cai et al., 2016). However, it has been challenging to retrieve large genome fragments or complete genomes from each AD complex microbial community members, due to insufficient sequencing depth and population heterogeneity (Morozova and Marra, 2008). Gene-centric methods, a specific focus on gene coding regions as opposed to non-coding regions, has been used to analyse unassembled data in complex environments such as AD, as they provide an overview of gene frequencies by mapping reads against protein databases (Schlüter et al., 2008).

Metagenomics was first applied to AD microbiomes in 2008. Microbial communities within a full-scale biogas plant, treating agricultural waste, were analysed and it was found that *Methanoculleus* played a dominant role in methanogenesis while *Clostridia* were essential for the hydrolysis of the feedstock (Schlüter et al., 2008). Since then, metagenomics has been used to study the
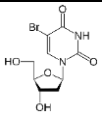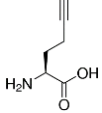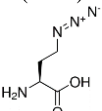
phylogenetic and functional diversity of multiple AD microbial communities from lab-scale reactors digesting co-substrates (Wirth et al., 2012; Zhu et al., 2019), animal manure (Li et al., 2013; Campanaro et al., 2016) and rice straw (Pore et al., 2016) through to full-scale reactors treating industrial waste (Wang et al., 2013; Cai et al., 2016; Chun-Te Lin et al., 2020), municipal waste (Wong et al., 2013; Yang et al., 2014), and activated sludge (Guo et al., 2015; Campanaro et al., 2018).

The combination of metagenomics and AD operational performance data enables us to investigate how functional capabilities of complex microbial communities carrying out the four focal processes (see section 1.3) within anaerobic digesters change over time. Nevertheless, metagenomics does not fully capture the plasticity of AD microbiomes and its effect on AD performance (Singer et al., 2017), because this technique is unable to discern between living, functionally active and dead cells or even extracellular relic DNA (Cangelosi et al., 2014; Lim et al., 2020).

1.5 Approaches for identifying in-situ active microbes in environmental samples

The effort to decipher the diversity, role and function of microbial communities and link these to their in-situ activity has become a "holy grail" (Urbach et al., 1999; Radajewski et al., 2000; Neufeld et al., 2007; Hatzenpichler et al., 2014; Hatzenpichler et al., 2016; Leizeaga et al., 2017; Couradeau et al., 2019). Researchers have applied several approaches that target different cell processes and coupled with shotgun nucleic acid sequencing for studying single-cell or mixed microbial communities' activity (Table 1.7; Figure 1.9). These approaches allow us to gain information on the active microbes present in environmental samples.

Table 1.7 Comparison of techniques for studying active microbial populations in environmental samples.

Tech.	Reporter	Cell process targeted	Advantage	Disadvantage	Assumptions	Analysis	Reference
DNA-/RNA-SIP	Stable heavy-isotope carbon (^{13}C)/ nitrogen (^{15}N)/ oxygen (^{18}O) / hydrogen (^{2}H) labelled substrate	DNA synthesis/ Transcription	Link cellular identity and function; High phylogenetic resolution	Long incubation period; cross-feeding issue; contamination risk; high GC DNA issue; reliance on commercially supplied labelled compounds	Substrate addition may increase cell division but not reflect in situ growth rates.	16s rRNA; Metagenomic	Neufeld et al., 2007; Bernard et al., 2007; Kalyuzhnaya et al., 2008; Chen et al., 2008; Neufeld et al., 2008; Sul et al., 2009; Winderl et al., 2010; Dumont et al., 2011; Pratscher et al., 2011; Gutierrez et al., 2013; Chemerys et al., 2014; Verastegui et al., 2014; Werner et al., 2014; Eyice et al., 2015; Fortunato and Huber, 2016; Singer et al., 2017; Ziels et al., 2018; Bradford et al., 2018; Gülay et al., 2019
BrdU	5-bromo-2'-deoxyuridine	DNA synthesis	When coupled with fluorescent antibody staining, it may be used as a single-cell technique to distinguish individual cells from the microbiome.	Low labelling efficiency	Rate of uptake varies by cell; may be toxic to some cells.	16s rRNA; Metagenomic	Mou et al., 2008; Edlund et al., 2008; David et al., 2015; Hamasaki et al., 2016
BONCAT	Homopropargylglycine (HPG) / Azidohomoalanine (AHA)	Translation	Link cellular identity and function; Fluorescence based in-situ activity study; Activity based cell sorting; Does not interfere with native protein synthesis or degradation; Short incubation period; Uses small amounts of materials/biomass	Methionine rich samples are tough	Cell growth may be stimulated by added amino acids; Dependence on uptake mechanism; Potential for cell inactivation or community shift	16s rRNA	Hatzenpichler et al., 2014; Samo et al., 2014; Hatzenpichler et al., 2016; Couradeau et al., 2019; Reichart et al., 2020; Steward et al., 2020
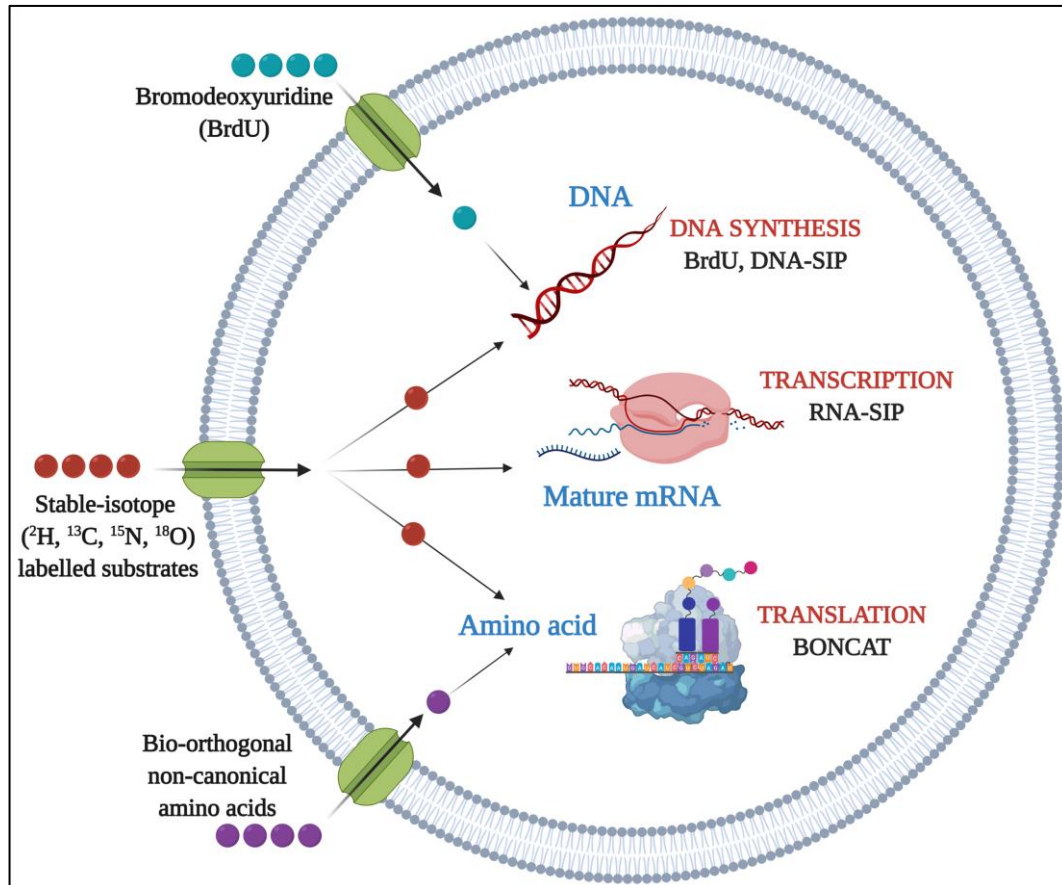
Figure 1.9 Approaches for labelling and targeting cell processes in an active microbial cell that can be coupled with (shotgun) nucleic acid sequencing. Adapted from Singer et al. (2017) and Hatzenpichler et al. (2020). Created with BioRender.com.

1.5.1 Bromodeoxyuridine (BrdU)

Bromodeoxyuridine (BrdU) labelling tags viable cells in-situ. This technique uses BrdU, a synthetic nucleoside analogue of thymidine, to label the nascent DNA of actively growing cells which can then be isolated using immunocapture techniques (Urbach et al., 1999). BrdU is added to a sample, incubated and then cells are collected by centrifugation. Cells that have taken up and incorporated BrdU into their newly synthesised DNA can then be isolated and BrdU-labeled DNA purified using an immunocapture technique (Borneman, 1999; Urbach et al., 1999). Subsequent sequencing of the BrdU-labelled DNA is used to identify

microbes that were actively growing in the sample. BrdU labelling has been applied to study active microbes that respond to specific carbon substrates in soil (Borneman, 1999; Yin et al., 2000; Goldfarb et al., 2011; David et al., 2015), sea sediment (Edlund et al., 2008), ocean water (Mou et al., 2008; Hamasaki et al., 2016) and lake water (Urbach et al., 1999).

1.5.2 Stable-isotope probing (SIP)

Stable-isotope probe (SIP) labelling is the most applied method to study in-situ microbial activity (Table 1.7; Singer et al., 2017). SIP is a technique that uses stable isotope (^{18}O, ^{13}C, or ^{15}N) labelled substrates to facilitate the selective enrichment of the DNA or RNA from active cells within complex environmental samples (Radajewski et al., 2000; Radajewski et al., 2003). Isotope-labelled, or "heavy", nucleic acids are isolated using density-gradient ultracentrifugation and subsequent gradient fractionation, and then purified using caesium chloride (CsCl) before sequencing analysis for active microorganisms' identification (Neufeld et al., 2007; Whiteley et al., 2007). This technique was first applied by Radajewski et al. (2000) to investigate in situ methanol-utilising microorganisms in soil by using various ^{13}C-enriched carbon source and eventually found active methylotrophs related to Alphaproteobacteria and Acidobacterium. DNA-SIP and RNA-SIP techniques allow the researcher to identify key players involved in the metabolism of certain compounds under in-situ conditions of soil (Bernard et al., 2007), aquifer sediment (Winderl et al., 2010), lake sediment (Dumont et al., 2011) and anaerobic digester sludge (Ziels et al., 2018) samples. However, this technique needed costly labelled substrates and required long incubation periods and a large amount of biomass (Dumont and Murrell, 2005).

1.5.3 Bioorthogonal non-canonical amino acid tagging (BONCAT)

In the past few years, researchers from several laboratories (Hatzenpichler et al., 2014; Samo et al., 2014; Hatzenpichler et al., 2016; Leizeaga et al., 2017; Couradeau et al., 2019) have applied a method called bioorthogonal non-canonical amino acid tagging (BONCAT) to study metabolically active microbes in pure culture and environmental samples, such as marine sediments, pond sediments, and soil. This technique uses a synthetic amino acid, i.e. L-azidohomoalanine (AHA) or L-homoproparglycine (HPG) as surrogates for L-methionine (Met) (Table 1.7; Kiick et al., 2002; Dieterich et al., 2006; Hatzenpichler et al., 2014). Any organism that is actively making and synthesising new proteins incorporates the synthetic amino acid into nascent proteins. There is an indication that this works in a broad range of microorganisms and does not interfere with native protein synthesis or degradation (Hatzenpichler et al., 2014; Samo et al., 2014). Thus, we are not limited to identifying the activity of one particular microorganism. The active cells can be visualized using fluorescence microscopy by attaching fluorescent dye via Cu(I)-click chemistry (Hatzenpichler et al., 2016) to the nascent protein that has incorporated the synthetic amino acid (Figure 1.10). Compared to other methods previously mentioned, BONCAT is rapid (the time required is less than 1 h) and highly sensitive, with BONCAT allowing detection of nascent proteins after only minutes of incubation using small amounts of materials/biomass (Couradeau et al., 2019). This makes it appealing for tracking activity or monitoring microbial community changes in anaerobic digestion.
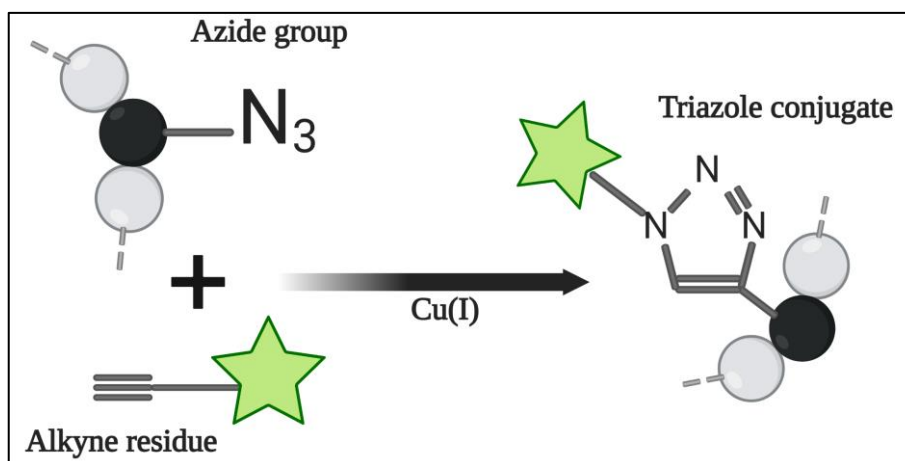
Figure 1.10 Cu(I)-catalysed click chemistry linking an azide labelled protein (-N$_3$) to a terminal alkyne residue of a fluorescent dye (star) to yield a triazole conjugate. Adapted from Hatzenpichler et al. (2014). Created with BioRender.com.

1.6 Bioinformatic approaches

Multiple bioinformatics methods and analytic approaches have been developed for the interpretation of large sequencing-based biological datasets. Traditionally metagenomic samples were analysed by using amplicon sequencing and using 16S rRNA gene sequences via short-read sequencing technology (e.g., Illumina) for species identification and abundance (Johnson et al., 2019). Short-read sequencing breaks DNA into small pieces that are amplified and subsequently sequenced to yield 'reads' (~150 – 300 bp). Genome reconstruction of short reads is computationally expensive and the generated assemblies are frequently very fragmented (Miller et al., 2010). Currently, genome reconstruction of AD communities has been greatly advanced by the development of long-read sequencing, e.g. Pacific Biosciences Single Molecule Real-Time (SMRT) and Oxford Nanopore Technologies MinION/GridION/PromethION. Long-read sequencing directly sequences single molecules of DNA in real time, typically without the need for amplification and produces longer reads (on average > 10 kbp).

Bioinformatic techniques are then used to piece together the long-reads like a jigsaw, into a continuous genomic sequence (Figure 1.11). Due to the longer read lengths fewer pieces need to be assembled and the resulting genomes are less fragmented (Schatz et al., 2010).
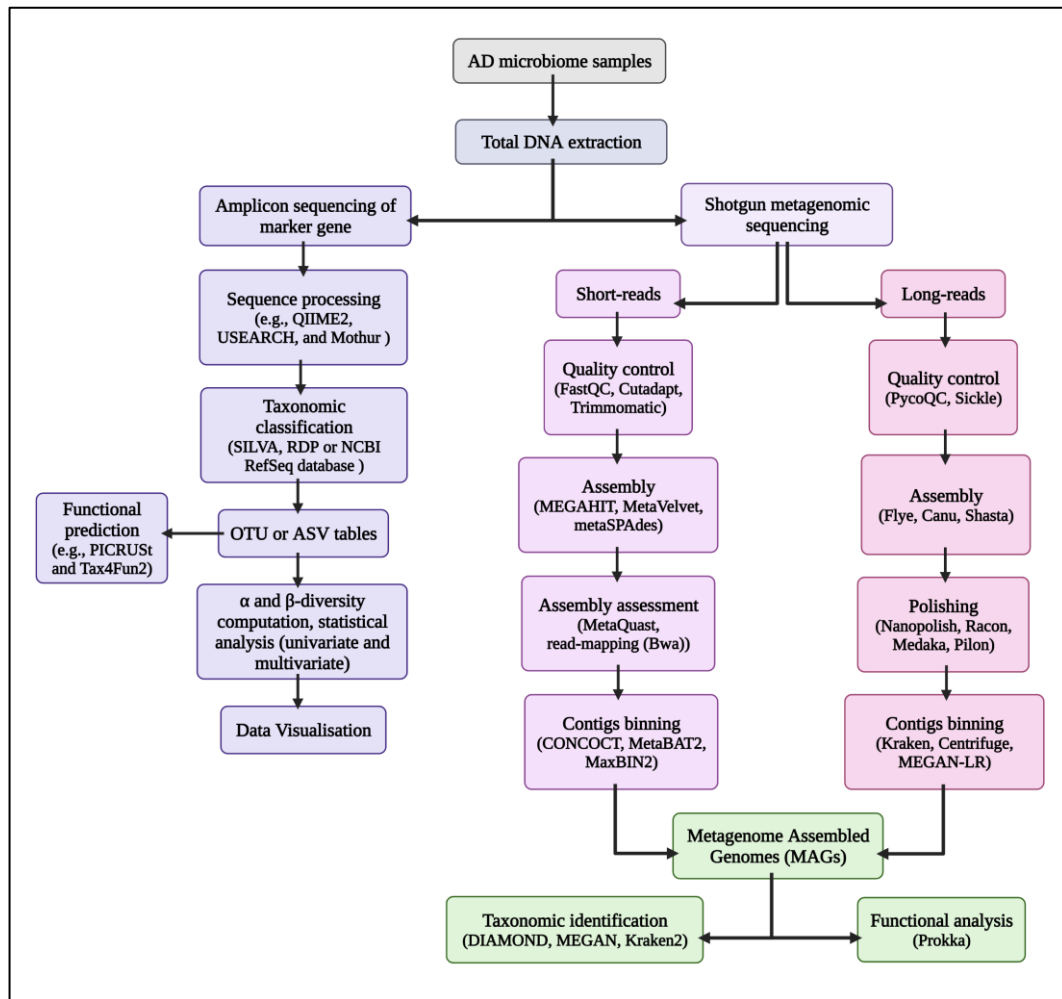
Figure 1.11 Example of bioinformatics pipelines for analysing marker gene and metagenomic dataset in AD microbiome studies. Created with BioRender.com

In marker gene-based approach, the whole 16S rRNA gene, specific regions, and/or the *mcrA* gene are amplified and barcoded by attaching a short unique sequence segment to label individual samples for multiplexing and simultaneous sequencing in a single sequencing run, particularly using the Illumina MiSeq platform. The sequencing reads are quality checked, processed, and

phylogenetically analysed for taxonomic classification by using reference gene databases (e.g. Greengenes, Silva, RDP and NCBI RefSeq). Several bioinformatic analysis pipelines have been developed and utilised for microbial marker gene analysis (Figure 1.11), such as QIIME2 (Bolyen et al., 2019), Bioconductor (Callahan et al., 2016), USEARCH (Edgar, 2010), Mothur (Schloss et al., 2009) and MG-RAST (Meyer et al., 2008). These pipelines enable us to calculate α diversity (Walters and Martiny, 2020), which defines the variety and structure of microbiomes in a sample, and β diversity (Walters and Martiny, 2020), which compares the similarity or dissimilarity of microbiomes from different samples. Afterwards, univariate and multivariate statistical analyses (e.g., principal component analysis, principal coordinates analysis, non-metric multidimensional scaling, partial least squares discriminant analysis, and Linear discriminant analysis Effect Size (LEfSe; Segata et al., 2011)) are utilised to compare AD microbiomes from different samples or to identify particular taxa that are influenced by the changes in AD operating conditions. As marker gene-based approaches can identify taxa but provide little information on their metabolic or functional capability, some bioinformatic tools have been developed (e.g. Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt; Langille et al., 2013), Tax4Fun2 (Wemheuer et al., 2020), Piphillin (Iwai et al., 2016), and PanFP (Jun et al., 2015)), to predict the genomes and functions of the identified taxa from the closely related known species. However, it should be emphasised that this type of functional inference is simply a prediction and does not offer a reliable functional profile.

High-throughput (short- or long-reads) sequencing of metagenomic DNA, quality control of reads, sequence assembly, gene prediction and annotation, and binning of contigs into metagenome-assembled genomes (MAGs) or genome bins are all part of the standardised metagenomics analysis pipeline (Figure 1.11). De novo assembly of metagenomic sequence reads is typically the most difficult phase since it is computationally intensive and time-consuming. To increase the speed and precision of de novo assembly, many algorithms and bioinformatic tools have been created. Assemblers such as MEGAHIT (Li et al., 2015), MetaVelvet (Namiki et al., 2012), metaSPAdes (Nurk et al., 2017), IDBA-UD (Peng et al., 2012), Genovo (Laserson et al., 2011) and SOAPdenovo2 (Luo et al., 2012) are utilised for short-read de novo assembly while Falcon (Chin et al., 2016), minimap2/miniasm (Li, 2016), metaFlye (Kolmogorov et al., 2020), Canu (Koren et al., 2017) and Shasta (Shafin et al., 2020) are used for long-, error-prone reads. The quality of assembled contigs should be assessed and polished by mapping sequence reads to contigs or comparing metagenome assemblies to close references using MetaQUAST (Mikheenko et al., 2016), Nanopolish (Loman et al., 2015), Racon (Vaser et al., 2017), Medaka (https://github.com/nanoporetech/medaka) and Pilon (Walker et al., 2014). Binning contigs into individual MAGs can be performed using program such as CONCOCT (Alneberg et al., 2014), MetaBAT2 (Kang et al., 2019) or MaxBIN2 (Wu et al., 2016).

Metagenomic functional annotation is performed by comparison to databases (e.g. Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), Clusters of Orthologous Groups of proteins (COGs), Integrated Microbial Genomes & Microbiomes (IMG/M), SEED and Pfam (Mistry et al., 2021)) using Prodigal

(Hyatt et al., 2010), Prokka (Seeman et al., 2014) or Kraken (Wood and Salzberg, 2014). However, it is important to understand that these annotations and database are constantly being updated and do not provide complete picture. This means that unfortunately there are many novel or uncharacterised taxa that will not have a meaningful functional annotation. Metabolic pathways can be reconstructed and/or modelled from the genes annotated using metabolic pathway databases, such as KEGG Pathway, MetaCyc Metabolic Pathway and BRENDA (Braunschweig ENzyme DAtabase) database. Metagenomic sequences can also be annotated by comparing them to "specialty databases," such as the CAZy database for carbohydrate-active enzymes and the MEROPS database for peptidases. Thus, metagenomics has the ability to recover genomes and predict the metabolism of novel and uncultured microbes.

1.7 Aims

Our understanding of microbial communities associated with anaerobic digestion (AD) currently relies strongly on marker gene profiling and metagenomic sequencing, which can reveal phylogenetic diversity but does not provide information concerning microbial activity or the close associations that may form between syntrophic organisms. To fill this gap, approaches that facilitate charting of process-targeted variation in microbial community activities are important for understanding how the microbiology of AD functions as a single biological machine. Therefore, the overarching goal of my PhD project is to adopt a top-down approach to narrow down the complex AD microbiome to subsets of metabolically active microbes as they respond to substrate availability in AD systems using bioorthogonal non-canonical amino acid tagging (BONCAT) and affinity-based

cell separation (ABCS) techniques. This combination of techniques could help to identify the key microbes involved in the degradation of compounds in AD systems, enhance our understanding of microbial community interactions, and facilitate the development of strategies for process optimisation. To do this, a number of approaches were taken:

1. To better understand the fate and intermediates generated during the catabolism of medium-chain fatty acids, individual VFAs with odd and even numbers of constituent carbon atoms were added to starved AD cultures. To apply BONCAT to track the metabolically active microbes, cells were exposed to the synthetic amino acid azidohomoalanine (AHA) to obtain a temporal snapshot of the active cell fraction, as described in Chapter 3.

2. To apply BONCAT to cell-selective analysis, AHA-labelled cells were enriched via reactions with affinity tags followed by downstream genomic analysis. Chapter 4 describes the application of temporally-selective cell labelling and enrichment using BONCAT and affinity-based cell separation (ABCS) in a simple system, using mixed pure cultures, for downstream genomic analyses to demonstrate the specificity and sensitivity of BONCAT-ABCS in a mixed system.

3. Chapter 5 describes the investigation of the active cell labelling and enrichment using BONCAT-ABCS in AD-derived sludge to identify which microbes actively contribute to specific metabolite processing in AD systems.

2 Materials and Methods

2.1 Inoculum

2.1.1 *Escherichia coli* strains and culture conditions

E. coli MV1300 (MG1655 Δ *lacZYA*; kanamycin (Kan) resistance) and *E. coli* MV1717 (MG1655 *lac+* and plasmid-encoded, inducible CDI-msfGfp, chloramphenicol (Cm) resistance), provided by Dr. Marjan van der Woude, were used for the mixed *E. coli* strains glucose-lactose diauxie experiment. Strains were grown at 37 °C overnight on Lysogenic Broth agar (LB-agar Miller; LMM0204, Formedium, Hunstanton, UK) containing 30 µg/mL kanamycin or 34 µg/mL chloramphenicol for MV1300 or MV1717, respectively. A single colony of each strain was grown overnight (18 hours) in 25 mL LB medium (LB broth, Miller; BP9723-500, Fisher BioReagents, Loughborough, UK) containing antibitotic (Kan 30 µg/mL or Cm 34 µg/mL) in 50 mL Falcon tubes at 37 °C with 120 rpm orbital shaking. The optical density at 600 nm (OD600) of each culture was measured (see section 2.3.5), the biomass from each tube was harvested via centrifugation (Centrifuge 5810 R, eppendorf) at $1940 \times g$, 37 °C. The supernatants were removed, the pellet resuspended in 10 mL of warm (37 °C) filter sterile 1x phosphate buffered saline (PBS 20-7400-10, Severn Biotech Ltd.). The biomass was spun again with the same parameters. The PBS was removed and 10 mL of 1x Morpholinepropanesulfonic acid (MOPS) minimal medium (Teknova, Hollister, CA, USA) was added into each tube. The strains were mixed together in 1:1 ratio (v/v) and OD600 measured. This mixed culture was then used as the inoculum for *E. coli* glucose-lactose diauxie experiments.

2.1.2 AD derived starved microbial community

'Starved' inocula were generated using material collected from a process-scale (1,858 m^3) AD system at Yorkshire Water's Naburn site, York, United Kingdom (53°54'50.5"N 1°05'04.6"W). Material was collected from an outlet pipe coming from digester tank 3 into a 30 L plastic barrel that was consequently sealed. The material was immediately transported to the lab where it was transferred to 5 L reactors for incubation.

Samples were incubated in stirred 5 L reactors (Figure 2.1) at 35 °C until VFA content was undetectable by GC-FID (HP 5890 series II) measurements (< 0.05 mM; see section 2.3.2). Starved communities were used as the starting point for experiments.

2.2 Bioreactor set up and operation

2.2.1 Applikon benchtop bioreactor operation

A glass autoclavable bioreactor with a 3 L working volume (Applikon Biotechnology, Delft, The Netherlands) was used for mixed *E. coli* strains diauxic growth experiments. The temperature was kept at 37 °C (± 0.3) via a water bath (OLS200, Grant Instruments). pH was logged and monitored by Bio Controller ADI 1010 (Applikon Biotechnology, Delft, The Netherlands) and maintained at 7.2 ± 0.05 by addition of 2 M NaOH. Dissolved oxygen was maintained above 20% saturation by adjusting agitation speed in the range of 270 – 500 rpm (Motor Controller, ADI 110, Applikon Biotechnology, Delft, The Netherlands) with fixed 1 L/min air flow. The mixed *E. coli* MV1300 and MV1717 strains (from section 2.1.1) were used to inoculate the 3 L fermentor with 1 L 1x MOPS minimal medium

(Teknova, Hollister, CA, USA) containing 0.5 g/L glucose and 1.5 g/L lactose as the only carbon source.

2.2.2 5 L anaerobic digesters operation

Lab scale, single stage anaerobic digesters with 5 L working volume were constantly mixed at approximately 45 rpm. Temperature was maintained at 35 $^{\circ}$C via a heating jacket and controlled via a custom feedback loop system. Spill overs and feeding tubes were closed in this experiment. Sample only from sampling tubes to help control gas volume measurements (Figure 2.1).
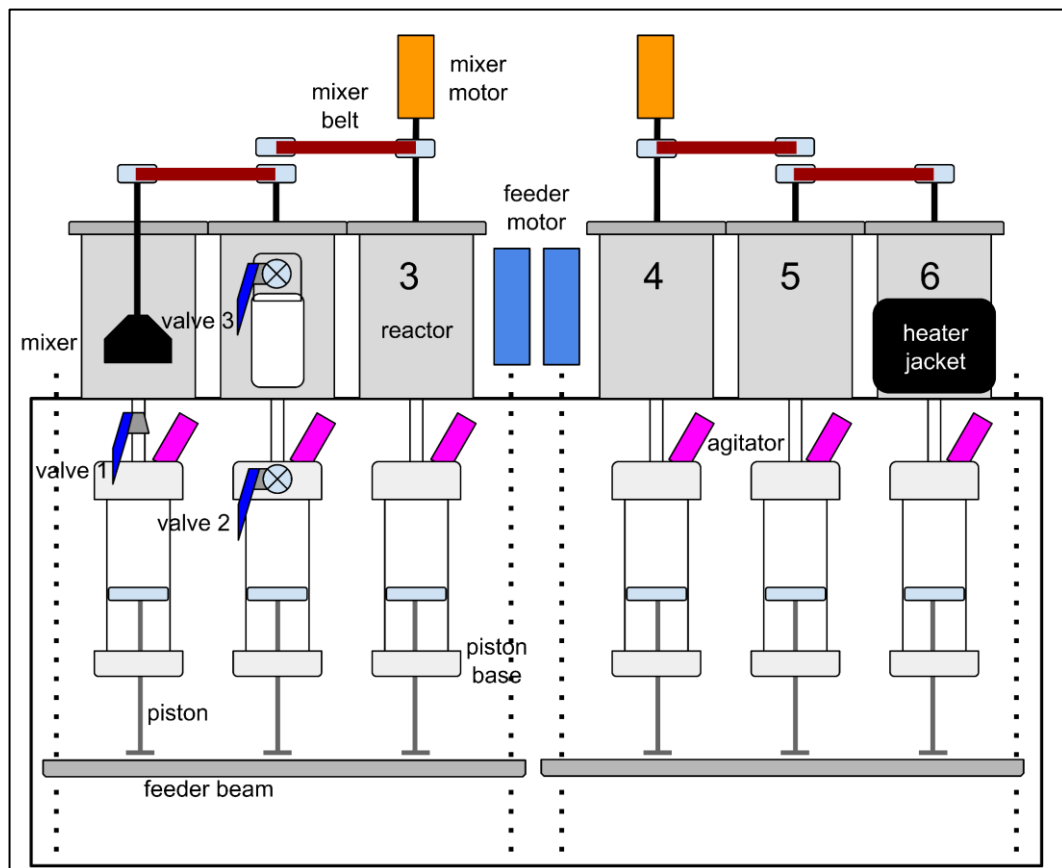
Figure 2.1 The system of 5 L reactors each with independent temperature controls and gas volume measurements. These are stirred at a fixed speed in banks of 3, they can also be fed at programmable intervals in banks of 3. The systems are controlled by two PLC systems. Credit: Prof. James Chong.

2.3 Process data

2.3.1 Volatile fatty acids (VFA)

5 mL samples were centrifuged (6,000 rpm, 15 min at 4 ^{o}C; Centrifuge 5810 R, eppendorf). Supernatants were 0.22 µm filtered (Millex) into 1.5 mL microfuge tubes (15562320, Fisherbrand, Germany). Filtrates were acidified (7.5 µL neat orthophosphoric acid (Scientific Laboratory Supplies CHE2710) / mL) before 1 µL was injected for VFA analysis. VFAs were detected on a GC-FID (HP 5890 series II) fitted with a Nukol capillary column (30 m × 0.25 mm, df 0.25 µm; 24107, Sigma). Helium carrier gas was flowed at a rate of 5 mL/min. Detectors and injectors were held at 200 ^{o}C and samples eluted via a temperature gradient of 100 – 150 (10 ^{o}C/min), 150 – 200 (20 ^{o}C/min), 200 °C hold for 10 minutes. The GC-FID was calibrated using Volatile Free Acid Mix (CRM46975, Sigma-Aldrich, Dorset, UK) providing C_2-C_7 reference values between 0 and 10 mM to generate 5 points calibration curves. Additional solutions of 0.1 – 10 mM octanoic (C8; 99% Sigma-Aldrich C2875, Dorset, UK), nonanoic (C9; 73982, Sigma-Aldrich, Dorset, UK) and decanoic acids (C10; 8021690100, Sigma-Aldrich, Dorset, UK) were used to extend the calibrated detection range. The calibration results showed a satisfactory correlation (R^2>99%) with limits of detection of 0.01 – 0.05 mM (Figure 2.2). DataApex Clarity^{TM} software was used for data acquisition, processing and instrument control.
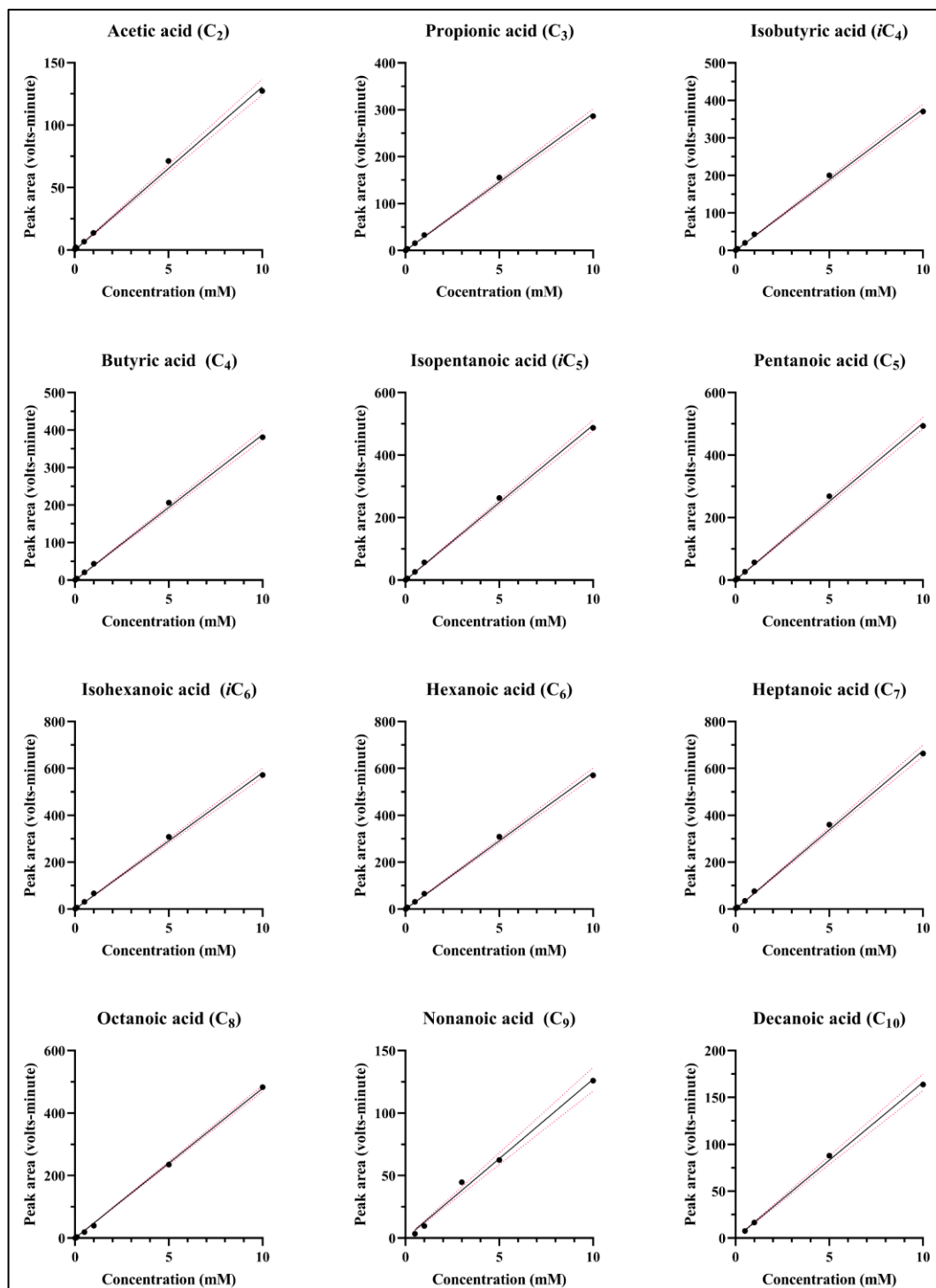
Figure 2.2 Calibration curves for each volatile fatty acid. Red-dot lines denote standard deviation (SD) of triplicates. The GC-FID was calibrated using Volatile Free Acid Mix providing C_2-C_7 reference values between 0 and 10 mM to generate 5 points calibration curves. Additional solutions of $0.1 – 10$ mM C_8, C_9 C_{10} were used to extend the calibrated detection range.

2.3.2 Glucose

Glucose concentration and standard curve (Figure 2.3) assays were performed using Glucose Assay Kit (MAK263, Sigma-Aldrich) following the manufacturer's instructions. The concentration was measured by assessing the optical density (OD570) using a microplate reader (CLARIOstar, BMG Labtech).

$$Y = 0.1223X + 0.04457$$
$$R^2 = 0.9919$$
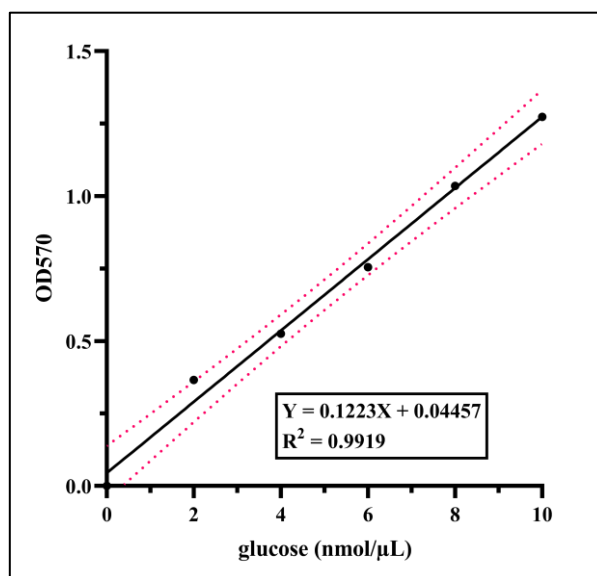
Figure 2.3 Glucose standard curve. Red-dot lines denote standard deviation (SD) of triplicates.

2.3.3 Lactose

Lactose concentration and standard curve (Figure 2.4) assays were performed using Lactose Colorimetric / Fluorometric Assay Kit (K624, BioVision) following the manufacturer's instructions. The concentration was measured by assessing the optical density (OD570) using a Thermo Spectronic Biomate 3 UV-Visible spectrophotometer (ThermoFisher Scientific, UK). 100 µL of the reaction mix was pipeted into a UV micro Cuvette (70 - 850 µL capacity; Z637092; Brand, Sigma-Aldrich) for reading. Reaction mix without lactose standard/sample was used as a blank. The amount of galactose measured is equal to lactose.
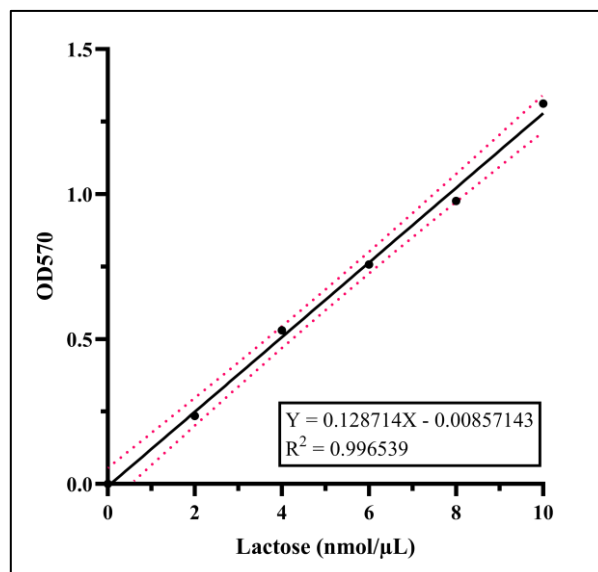
Figure 2.4 Lactose standard curve. Red-dot lines denote standard deviation (SD) of triplicates.

2.3.4 Cell density (OD600)

E. coli growth in LB or 1x MOPS minimal medium was measured by optical density (OD600) using a Thermo Spectronic Biomate 3 UV-Visible spectrophotometer (ThermoFisher Scientific, UK). 500 µL culture was measured in a semi-micro cuvette (1.5 mL capacity; Kartell^{TM} 0193800, Fisher Scientific, UK). Cultures were diluted in LB or 1x MOPS minimal medium and remeasured when OD600 > 0.4. Sterile LB or 1x MOPS minimal medium was used as a blank.

2.4 BONCAT labelling and enrichment

2.4.1 Click labelling of chemically fixed microbial cells

Sample preparation for on-slide and in-solution click labelling were essentially as described (Hatzenpichler et al., 2014; Hatzenpichler and Orphan, 2015) using the Click-&-Go^{TM} Cell Reaction Buffer Kit (Click Chemistry Tools, Scottsdale, AZ, USA). Freshly mixed cocktail was prepared following the manufacturer's instructions (per 1 reaction): 440 µL of 1× reaction buffer, 10 µL Copper(II)

Sulfate, 50 µL reducing agent. Click cocktails were supplemented to a final concentration of 3 µM Alexa-488 alkyne or Biotin alkyne (Click Chemistry Tools, Scottsdale, AZ) for visualization or cell sorting samples, respectively. Cocktail solutions were used within 10 min of preparation.

2.4.2   BONCAT visualisation

Fixed AHA-labelled biomass was diluted 10x in ultrapure water and spread on Teflon-coated glass slides (ER-208B-CE24, Thermo Scientific). Slide preparations were dried (46 ºC, 3 min). Cells were dehydrated and permeabilized by sequentially immersing slides in 50%, 80%, and 96% ethanol in ultrapure water for 3 min before being air-dried at room temperature. 20 µL click reaction cocktail containing Alexa-488 alkyne was applied to a cleaned coverslip, inverted and placed over the fixed sample (Leizeaga et al., 2017). Prepared slides were incubated in a dark humid chamber (RT, ≥30 min) (Hatzenpichler and Orphan, 2015), then washed three times with PBS (3 min). Afterwards, slides were subjected again to the dehydration protocol described above before being air-dried at RT. Samples were mounted with VECTASHIELD® Mounting Medium with DAPI (H-1200, Vector Laboratories Ltd, Peterborough, UK) and examined by confocal microscopy (Zeiss LSM 710). Images were captured and analysed using ZEN imaging software (Zeiss). Samples were identified via DAPI before switching to other fluorescent channels to avoid bias for BONCAT positive microbial aggregates (Hatzenpichler et al., 2014). Representative images were taken for each sample. Ten images of each samples were captured randomly to calculate the percentage of metabolically active cells (BONCAT-stained cells) that related to the DAPI-stained cells according to Chen et al. (2021).

2.4.3 BONCAT-labelled biomass recovery via Affinity-based cell separation (ABCS)

250 µL of fixed AHA-labelled samples were pelleted by centrifugation (16,100 × g, 5 min, RT; Sigma Sciquip 1-14). Pellets were resuspended in 250 µL of 80% ethanol and incubated (3 min, RT). Resuspended cells were supplemented with 1.2 mL of 96% ethanol, mixed and incubated (3 min, RT) before pelleting, then washed in 250 µL 1x phosphate buffered saline (PBS 20-7400-10, Severn Biotech Ltd.) and resuspended in 500 µL click reaction cocktail containing Biotin alkyne by vortexing (3 sec). The reaction was mixed by rotation (30 min, RT). Cells were washed with 800 µL PBS then resuspended in 800 µL (20% glycerol in 1x PBS) by vortexing (5 sec). 10 µL Neutravidin magnetic beads (78152104010150, GE Healthcare) were added to each sample, and mixed by rotation (30 min, 4 °C). Beads were concentrated on a magnetic stand (5 min) before discarding the supernatant. BONCAT-labelled / separated biomass was used for DNA extraction.

2.4.4 BONCAT labelled-protein enrichment

800 mg of cell pellets were resuspended in lysis buffer (8 M urea, 200 mM Tris pH 8, 4 % CHAPS, 1 M NaCl; Click Chemistry Tools), and treated with protease inhibitor (cOmplete, Mini, EDTA-free Protease Inhibitor Tablet; Roche). To fully lyse the cells, lysates were put on ice for 10 min, sonicated with a microtip probe by applied twelve 3 sec pulses with an amplitude of 32 – 33 % and pulse 3 sec on/5 s off (Bandelin Sonopuls HD 2070). The lysis mixture was incubated on ice for 1 minute after a cycle of three 3 sec pulses. Lysates were clarified by centrifuging at 10000 × g for 5 min.

For enrichment, approximately 200 µL of washed Dibenzocyclooctyne-agarose resin (50 % slurry; Click chemistry Tools) and 1 mL of 2x Catalyst solution (Click chemistry Tools) were added to 800 µL of each lysate. A freshly made 2× catalyst solution per enrichment consisted of 860 µL 18 MΩ water, 100 µL reaction additive 1 (Component D), 20 µL Copper (II) Sulfate solution (Component E) and 20 µL reaction additive 2 (Component F) (1033, Click-&-Go Protein Enrichment Kit, Click Chemistry Tools, Scottsdale, AZ, USA). Resin-treated lysates were rotated end-over-end at RT for 20 h.

To reduce and alkylate the resin bound protein, resin samples were washed with 18 MΩ water, treated with 1 mL of SDS wash buffer (100 mM Tris, 1 % SDS, 250 mM NaCl, 5 mM EDTA, pH 8; Click Chemistry Tools) and 10 µL 1 mM dithioethritol (DTT) for 15 min at 70 °C then cooled at RT for 30 min. Resin samples were subsequently treated with 1 mL 40 mM iodoacetamide in the dark for 30 min at RT.

Resin samples were transferred to a Spin Column (Component H; Click Chemistry Tools) and underwent stringent washing with 5×2 mL of SDS Washing Buffer (Component G; Click Chemistry Tools). The resin was subsequently washed with 10×2 mL 8 M urea in 100 mM Tris pH 8 and 10×2 mL 20 % Acetonitrile in water (v/v) to achieve additional stringent removal of non-specifically bound protein and SDS prior to on-resin digestion and mass spectrometry analysis of the enriched proteins. Resin samples were resuspended in Digestion Buffer (100 mM Tris, 2 mM $CaCl_2$, 10% acetonitrile) and pelleted by centrifugation for 5 min at $1000 \times$ g. Supernatants were removed to yield approximately 200 µL of digestion buffer in the tube with the resin. Samples were then sent to The York Centre of

Excellence in Mass Spectrometry, University of York CoEMS for mass spectrometry analysis (see sec. 2.5.4).

2.5   Molecular methods for community analysis

2.5.1   Genomic DNA extraction and quantification

2.5.1.1   Qiagen DNeasy Blood and Tissue Kit

Genomic DNA of *E. coli* strains was extracted using a Qiagen DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol with the following exceptions: samples (maximum 5×10^6 cells) were centrifuged for 10 min at $300 \times$ g; after addition of Buffer AE, incubated for 2 minutes at room temperature. The elution step was repeated 4 times to increase DNA yield.

2.5.1.2   Qiagen Powersoil DNA extraction

Genomic DNA was extracted from BONCAT-labelled AD biomass using a Qiagen PowerSoil DNA Extraction Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol with the following exceptions: samples were bead beaten for 20 min; after addition of ethanol wash solution (C5) samples were centrifuged for 60 sec.

2.5.1.3   DNA quantification

DNA concentrations were quantified using a Qubit^{TM} 1x dsDNA HS Assay Kit (Q33230, Invitrogen, Thermo Fisher Scientific) and Qubit^{TM} 3.0 Fluorometer (Q33216, Invitrogen, Thermo Fisher Scientific). Samples were stored at -80 °C.

2.5.2   PCR and qPCR

2.5.2.1   Primer design

Primers were designed to detect 3 different genes (Table 2.1). All primers were synthesised by Sigma-Aldrich with normal desalting purification. Dried primers were resuspended in 1x TE (10 mM Tris, pH 7.5 – 8.0, 1 mM EDTA). Primers were stored at – 20 °C. End point PCR was carried out with each primer pair to ensure a single band was formed for each selected target (see Figure 4.4).

Table 2.1 Primer sequences used to amplify the housekeeping gene (*rpsQ*; this study) and the target gene, kanamycin resistance (*aph(3')-II nptII*; Ullmann et al., 2019) and chloramphenicol resistance (*catA1*; Pholwat et al., 2019) gene.

Gene	Primer	Sequence (5 → 3)	Annealing (Tm)	GC %	Expected amplicon (bp)
rpsQ	Forward	GCACGTACATGACGAGAACA	62	50	97
	Reverse	AACCAGCGTCCAGGATTTAG	62	50	
nptII	Forward	GATCTCCTGTCATCTCACCTTGCT	61.96	50	129
	Reverse	TCGCTCGATGCGATGTTTC	58.71	52.6	
catA1	Forward	GCCAATCCCTGGGTGAGTTT	60.25	55	110
	Reverse	ACCTTGTCGCCTTGCGTATAA	60.07	47.6	

2.5.2.2   PCR and gel electrophoresis

The total volume for each PCR reaction was 50 µL, which consisted of 1 ng – 1 µg template DNA, 0.2 µM forward primer, 0.2 µM reverse primer, 25 µL One*Taq* Quick-Load 2x Master Mix with Standard Buffer (M0486, New England Biolabs)

and made up to 50 µL using nuclease-free water. PCR amplification was carried out using a Prime Thermal Cyclers (5PRIMEG/02; Techne Prime), where initial denaturation was at 94 °C for 30 seconds, followed by 94 °C for 30 seconds, 52 °C for 30 seconds and 68 °C for 15 seconds for 30 cycles and a final extension at 68 °C for 5 minutes. The PCR products were run on a 3 % agarose gel, using 3 g agarose (Agarose MB1200, Melford, Ipswich, UK), 100 mL 1x Tris-Borate EDTA (TBE) buffer and 3 µL SYBRSafe DNA gel stain (S33102, Invitrogen, UK). 20 µL sample was loaded into each well, and run in 1x TBE for 60 minutes at 100 Volts (5 V/cm).

2.5.2.3   qPCR

qPCR amplification and analysis were performed using the Applied Biosystems® QuantStudio® 3 Real-Time PCR System (A28567, Thermo Fisher Scientific). The qPCR mixture of 20 µl was prepared using 10 µl of FAST SYBR 2x Master Mix (Thermo Fisher Scientific), 0.7 µl of each primer (final concentration 350 nM), 6.6 µl of nuclease free water, and 2 µl template DNA (0.09375 – 6.02 ng/µL).

The thermal cycling protocol was as follows: initial denaturation for 20 s at 95 °C followed by 40 cycles of 1 s at 95 °C and 20 s at 60 °C. The fluorescence signal was measured at the end of each extension step at 60 °C. After amplification, a melting curve analysis with a temperature gradient of 0.1 °C/s from 60 to 95 °C was performed to confirm that only specific products were amplified. PCR efficiencies were calculated from the given slopes in QuantStudio® 3 software. An absolute

quantification of the *catA1* and *nptII* genes normalised to the housekeeping gene (*rpsQ*) was performed.

2.5.3 Metagenomic sequencing

2.5.3.1 Oxford Nanopore metagenomic sequencing

DNA sequencing was performed by the Bioscience Technology Facility (University of York, UK) using Oxford Nanopore Technologies' (ONT) MinION sequencing platform. Isolated DNA samples were purified and small fragments removed using AMPure XP beads (Beckman Coulter) at a 0.8:1 (bead:sample) ratio, with extended incubations for bead binding and elution. Library preparation was performed using ONT's ligation sequencing kit (SQK-LSK109) with barcoding expansion pack (EXP-NBD104). Individual samples were pooled into a single library for sequencing, using ONT's recommended protocol, modified as below. All additional enzymes required were supplied by New England Biolabs. 900 ng DNA per sample was prepared for barcode ligation using a combined FFPE nick repair and end repair step (30 min incubations at 37°C then 60°C). Following a clean-up step using AMPure XP beads, unique ONT barcode sequences for each sample were ligated with NEB Blunt/TA ligase mastermix (45 min incubation). Free barcodes were removed by AMPure bead clean up, and samples were quantified using a Qubit 3.0 fluorimeter with High Sensitivity dsDNA reagent (Invitrogen). Approximately 100 ng of each barcoded sample were pooled and ONT adapter (AMII) was added in a ligation reaction using NEBNext Quick T4 DNA ligase (30 min, RT). A final clean-up was performed with 0.5:1 (AMPure XP beads:sample), including two washes with ONT long fragment wash buffer, before

eluting the final library into 15 µl elution buffer. The library was loaded onto a MinION flow cell (FLO-MIN106, R9.4.1), and sequenced (48 h), with regular top-ups of flush buffer plus tether from the EXP-FLO002 flow cell priming kit. High accuracy base-calling and barcode demultiplexing was performed using ONT Guppy base-calling software version 3.2.4.

2.5.3.2 Illumina HiSeq metagenomic sequencing

DNA samples were sent to Novogene to be sequenced using three lanes on the Illumina HiSeq 2500 sequencing platform. The libraries were prepared using the NEBNext Ultra DNA Library Prep Kit for Illumina (E7370, NEB), according to the manufacturer's instructions. DNA samples were sheared to an average size of 200 bp using a Covaris S2 system before Library Prep Kit was used. After the PCR amplification step, fragment sizes were determined using the Agilent Bioanalyzer with Agilent High Sensitivity DNA Kit. The prepared libraries were run using a HiSeq2500 PE flow cell on the Illumina HiSeq 2500 platform.

2.5.4 LC-MS/MS

Liquid chromatography-tandem-mass spectrometry (LC-MS/MS) experiments were carried out by CoEMS. Protein samples were on-bead digested with 10 µL 0.1 µg/µL Promega sequencing grade trypsin and incubated at 37 °C overnight. Resulting peptides were desalted with Millipore C18 ZipTip before being re-suspended in aqueous 0.1% trifluoroacetic acid (v/v) then loaded onto an mClass nanoflow UPLC system (Waters) equipped with a nanoEaze M/Z Symmetry 100 Å C_{18}, 5 µm trap column (180 µm x 20 mm, Waters) and a PepMap, 2 µm, 100 Å, C_{18} EasyNano nanocapillary column (75 µm x 500 mm, Thermo). The trap wash

solvent was aqueous 0.05% (v:v) trifluoroacetic acid and the trapping flow rate was 15 µL/min. The trap was washed for 5 min before switching flow to the capillary column. Separation was achieved using an elution gradient of two solvents: solvent A, aqueous 0.1% (v:v) formic acid; solvent B, acetonitrile containing 0.1% (v:v) formic acid. The flow rate for the capillary column was 300 nL/min and the column temperature was 40°C. The linear multi-step gradient profile was: 3-10% B over 7 mins, 10-35% B over 30 mins, 35-99% B over 5 mins and then proceeded to wash with 99% solvent B for 4 min. The column was returned to initial conditions and re-equilibrated for 15 min before subsequent injections.

The nanoLC system was interfaced with an Orbitrap Fusion Tribrid mass spectrometer (Thermo) with an EasyNano ionisation source (Thermo). Positive ESI-MS and MS^2 spectra were acquired using Xcalibur software (version 4.0, Thermo). Instrument source settings were: ion spray voltage, 1,900 V; sweep gas, 0 Arb; ion transfer tube temperature; 275°C. MS^1 spectra were acquired in the Orbitrap with: 120,000 resolution, scan range: *m/z* 375-1,500; AGC target, 4e^5; max fill time, 100 ms. Data dependant acquisition was performed in top speed mode using a 1 s cycle, selecting the most intense precursors with charge states >1. Easy-IC was used for internal calibration. Dynamic exclusion was performed for 50 s post precursor selection and a minimum threshold for fragmentation was set at 5e^3. MS^2 spectra were acquired in the linear ion trap with: scan rate, turbo; quadrupole isolation, 1.6 *m/z*; activation type, HCD; activation energy: 32%; AGC target, 5e^3; first mass, 110 *m/z*; max fill time, 100 ms. Acquisitions were arranged by Xcalibur to inject ions for all available parallelizable time.

2.6   Bioinformatics

2.6.1   Contig Assembly and Polishing

Nanopore long-reads were assembled using Flye version 2.8 (Kolmogorov et al.,
2020) with the following parameters: metagenome mode for assembling uneven
coverage data (--meta), using raw nanopore reads (--nano-raw), 3 iterations of
polishing (--iterations 3) and an estimated genome size of 500m (--genome-size
500m). The resulting contigs were polished using the long reads with one round of
Medaka   version   0.11.3   using   consensus   mode
(https://github.com/nanoporetech/medaka). The base accuracy of this assembly was
then improved by using three cycles of Pilon version 1.23-java1.8 (Walker et al.,
2014) to polish the assembly using Illumina reads. These Illumina reads were
mapped to the assembly to generated a BAM file prior to Pilon using BWA version
0.7.17 and the mem algorithm (Li, 2013), and using Samtools version 1.9 (Danecek
et   al.,   2021)   for   indexing.   An   in-house   custom   pipeline
(https://github.com/ac1513/CLUSTard) was used to map raw reads to individual
contigs and visualise changes in abundance between samples. This was run using a
Pearson's correlation to discriminate between clusters designated as metagenome
assembled genomes (MAGs). It also used software such as checkM to evaluate the
validity of MAGs derived from this pipeline, generate a kraken2 and gtdb taxonomy
annotation and visualise abundance changes between sample treatments.

2.6.2   Prokka

Open Reading Frames (ORFs) from the MAGs were assigned using Prokka
(Seemann, 2014), using the default settings. These annotations were generated as

part of the CLUSTard pipeline, and amino acid predictions were used to generate a database to query the proteomic matches against.

2.7   Proteomic data analysis

Tandem mass spectra were extracted and charge state deconvolution and deisotoping were not performed. All MS/MS samples were analysed using Mascot (Matrix Science, London, UK; version 2.7.0.1). Mascot was set up to search against BONCAT, non-BONCAT and contaminant database (the non_boncat_20210921.fasta; 20120229c.fasta; PROKKA_D474_03102021 database (unknown version, 872063 entries)) assuming the digestion enzyme trypsin. Mascot was searched with a fragment ion mass tolerance of 0.50 Da and a parent ion tolerance of 3.0 PPM. O-124 of pyrrolysine, j-16 of leucine/isoleucine indecision and carbamidomethyl of cysteine were specified in Mascot as fixed modifications. Oxidation of methionine was specified in Mascot as a variable modification.

Scaffold (version Scaffold_5.1.0, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 22.0% probability to achieve an FDR less than 5.0% by the Percolator posterior error probability calculation (Käll et al., 2008). Protein identifications were accepted if they could be established at greater than 6.0% probability to achieve an FDR less than 5.0% and contained at least two identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii, 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis

alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters.

Only protein hits that passed the aforementioned filtering were used in the following analysis. Prokka accessions derived from the assembly annotation were used in peptide searches of proteomic hits against a database populated by prokka annotations derived from MAGs identified in the metagenomic data in mascot. Matches between metagenomic and proteomic data were made with a blastn database from BLAST 2.2.31+, which included these prokka accessions, and an evalue cutoff of 0.001. Other blast parameters remained default. Only MAG hits and proteomic hits matching back to the sample condition were selected from these matches, and the longest, highest scoring hit per MAG- proteomic match was used. This was done using sorting by query name, descending order of bit score, and ascending order of e value, and then taking the first result of these sorted queries per query name. Total spectrum counts were used for peptide abundance. Other level annotations and KEGG information are in Appendix H and I.

2.8 Software and database use

Additional software used included the following. Graph plots, bar plots and boxplots were generated using Prism 9 (v 9.0.2). Abundance graphs were generated by R (v 4.1.1). The chemical structures and reactions were drawn in Biorender.com. The Genome Taxonomy Database (GTDB; Release 06-RS202) was used for taxonomic classification of assigned metagenome-assembled genomes (MAGs).

3 Monitoring AD community fatty acid catabolism dynamics and tracking of metabolically active microbes in AD samples

3.1 Introduction

Anaerobic co-digestion of energy-dense waste such as fat, oil, and grease (FOG) has recently gained a great deal of attention for enhancing biomethane recovery in waste water treatment plants (Ziels et al., 2016). Methane yield has been reported to increase by 137 – 317% using this approach (Wan et al., 2011; Wang et al., 2013; Kurade et al., 2019). FOG induces higher biogas yield due to its higher potential degradable fraction (94.8%) than that of carbohydrates (50.4%) and proteins (71%; Jeganathan et al., 2006).

In AD, FOG is rapidly hydrolysed to yield free fatty acids (FAs), which are organic molecules composed of a hydrophilic head, a carboxyl group, and a varying length of hydrophobic aliphatic tail (Heukelekian and Mueller, 1958). FAs are primarily degraded by fermentative bacteria via β-oxidation, which results in a FA molecule with two fewer carbons, a molecule of acetyl-CoA, and highly reduced chemical compounds, *e.g.*, $FADH_2$ and $NADH_2$ (Jimenez-Diaz et al., 2017). To proceed β-oxidation to the next cycle, the supply of electron carriers (FAD^+ and NAD^+) are regenerated by a soluble hydrogenase that catalyses the oxidation of the highly reduced compounds and releases molecular hydrogen (H_2) or by a cytoplasmic formate dehydrogenase (FDH) that act as a CO_2 reductase to form formic acid (Agne et al., 2021). Hydrogen (H_2) and formic acid produced during β-oxidation are consumed by hydrogenotrophic methanogens to form methane.

Acetic acid may be metabolised to methane by acetoclastic methanogens in the absence of other electron acceptors, such as sulphate (Lalman and Bagley, 2000). Thus, anaerobic degradation of fatty acids is mostly the result of complex syntrophic partnerships between fermenting bacteria and methanogenic archaea (Worm et al., 2014).

So far, most experiments on the degradation of fatty acids in AD are based on the effect of long-chain fatty acids (LCFAs; $C_{14} – C_{20}$) on community composition and function (Angelidaki and Ahring, 1992; Neves et al., 2009; Sousa et al., 2009; Zhang et al., 2011; Ziels et al., 2016). In addition, these experiments merely compare the substrate and the β-oxidation final product (acetate) and report less information about potential intermediates formed during the course of β-oxidation. However, understanding the fate and function of intermediates generated from β-oxidation of FAs is critical in explaining the co-existence of active microbial species involved in fatty acid catabolism in AD. This chapter describes catabolism of medium-chain fatty acids and intermediates generated during the sequential removal of 2-carbon acetate groups by the AD microbial community. Fatty acid catabolism dynamics are monitored using a targeted gas chromatography (GC) method that has been the method of choice in fatty acid analysis for more than half a century (Seppänen-Laakso et al., 2002).

Furthermore, it will detail the use of bio-orthogonal amino acids coupled with click chemistry, called the bio-orthogonal non-canonical amino acid tagging (BONCAT) method, for obtaining a temporal snapshot of the active cell fraction during fatty acid catabolism. BONCAT relies on cellular uptake and incorporation of non-canonical amino acid, e.g., *L*-azidohomoalanine (AHA; the *L*-methionine

(Met) surrogate bearing a chemically-modifiable azide group) into nascent protein by exploiting methionyl-tRNA synthetase (MetRS) substrate promiscuity (Kiick et al., 2002). MET activation rate by MetRS is k_{cat}/K_M 5.47×10^{-1} $s^{-1}.\mu M^{-1}$, meanwhile, AHA has the highest activation rate (1.42×10^{-3} $s^{-1}.\mu M^{-1}$) amongst other methionine analogues, e.g., homopropargylglycine (HPG; 1.16×10^{-3} $s^{-1}.\mu M^{-1}$) and norleucine (5.22×10^{-4} $s^{-1}.\mu M^{-1}$). Accordingly, AHA has the highest incorporation efficiencies during *de novo* protein synthesis compared to HPG and norleucine (Kiick and Tirrell, 2000; Kiick et al., 2001; Kiick et al., 2002). AHA is water-soluble (max concentration 18.06 mg/mL), nontoxic, stable under most physiological and environmentally relevant conditions (except highly sulfidic and alkaline environments (Hatzenpichler et al., 2014)) and non-disruptive on protein synthesis and degradation in incubation with up to 1 mM AHA (Dieterich et al., 2006; Bagert et al., 2014; Hatzenpichler et al., 2014). Labelled proteins from active cells can be visualized using fluorescence microscopy after attaching fluorescent dyes with terminal alkyne via Cu(I)-catalyzed click chemistry (Huisgen, 1963; Rostovstev et al., 2002; Meldal et al., 2002) to the incorporated azide-bearing amino acids. BONCAT using AHA has high specificity as only one organism (*Karenia brevis*, a marine dinoflagellate formerly known as *Gymnodinium breve* or *Ptychodiscus breve*, which forms toxic red tide blooms along the Florida coast and the Gulf of Mexico) is currently known to produce this azide-containing metabolite (Griffin, 1994).

This method has been applied to study metabolically active microbes in pure culture and environmental samples, such as marine sediments, pond sediments, and soil (Hatzenpichler et al., 2014; Samo et al., 2014; Hatzenpichler et al., 2016;

Leizeaga et al., 2017; Couradeau et al., 2019), making it appealing for tracking the activity or monitoring microbial community changes in anaerobic digestion.

3.2 Experimental Design

The experimental approach used in this study was designed to track anaerobic digestion fatty acid catabolism products from even- and odd-numbered medium-chain fatty acids (Fig. 3.1A). 'Starved' inocula were generated using material collected from a process-scale (1,858 m^3) AD system at Yorkshire Water's Naburn site, York, United Kingdom (53°54'50.5"N 1°05'04.6"W) as described in section 2.1.2. Samples were incubated in gas-tight bottles at 35 °C until VFA content was undetectable by GC-FID measurements (< 0.05 mM). Throughout the starvation period, excess gas generated by community activity was vented periodically to relieve headspace pressure. Starved communities were used as the starting point for experiments.

After flushing the headspace with N_2 for 5 minutes, incubations of 50 mL starved samples were supplemented with single addition 'spikes' of 143 µL 6.981 M heptanoic (~20 mM C_7 final concentration), 79 µL 6.310 M octanoic (~10 mM C_8 final concentration), 143 µL 5.668 M nonanoic (~20 mM C_9 final concentration), or 98 µL 5.080 M decanoic acid (~10 mM C_{10} final concentration). Incubations of supplemented samples were performed in triplicate in 100 mL serum bottles stoppered with butyl rubber bungs at 35 °C. 2 mL samples were collected for VFA analysis (see sec. 2.3.1).

To track the active microbes during fatty acid degradation (Fig. 3.1B), a single dose of 10 mM octanoic acid (C_8) was added and incubated at 35 °C for 24 h.

Subsampling for VFA analysis and click chemistry-mediated fluorescence labelling (Hatzenpichler and Orphan, 2015) was conducted every 6 h. To visualize metabolically active microbes, 3 mL sample from each timepoint were transferred anaerobically into anaerobic glass tubes that had been flushed with pure nitrogen for 3 minutes. Each sample was supplemented with 1 mM final concentration of L-azidohomoalanine (AHA; Click Chemistry Tools, Scottsdale, AZ, USA) in nano-pure water (filter sterilize using 0.2 µm filter, pH 7) and incubated for 30 minutes at 35 °C. At the end of the incubation period, samples were centrifuged at 6,000 rpm for 15 min (Eppendorf Centrifuge 5810R) and the supernatant was subjected to VFA analysis (see sec. 2.3.1). The pellet containing AHA-tagged cells was fixed with ethanol:PBS (1:1 v/v) and stored at -20 °C. Sample preparation for click labelling of chemically fixed microbial cells and BONCAT visualisation was essentially as described in section 2.4.2. Ten images of each sample were captured randomly to calculate the percentage of metabolically active cells (BONCAT-stained cells) that related to the DAPI-stained cells according to Chen et al. (2021).
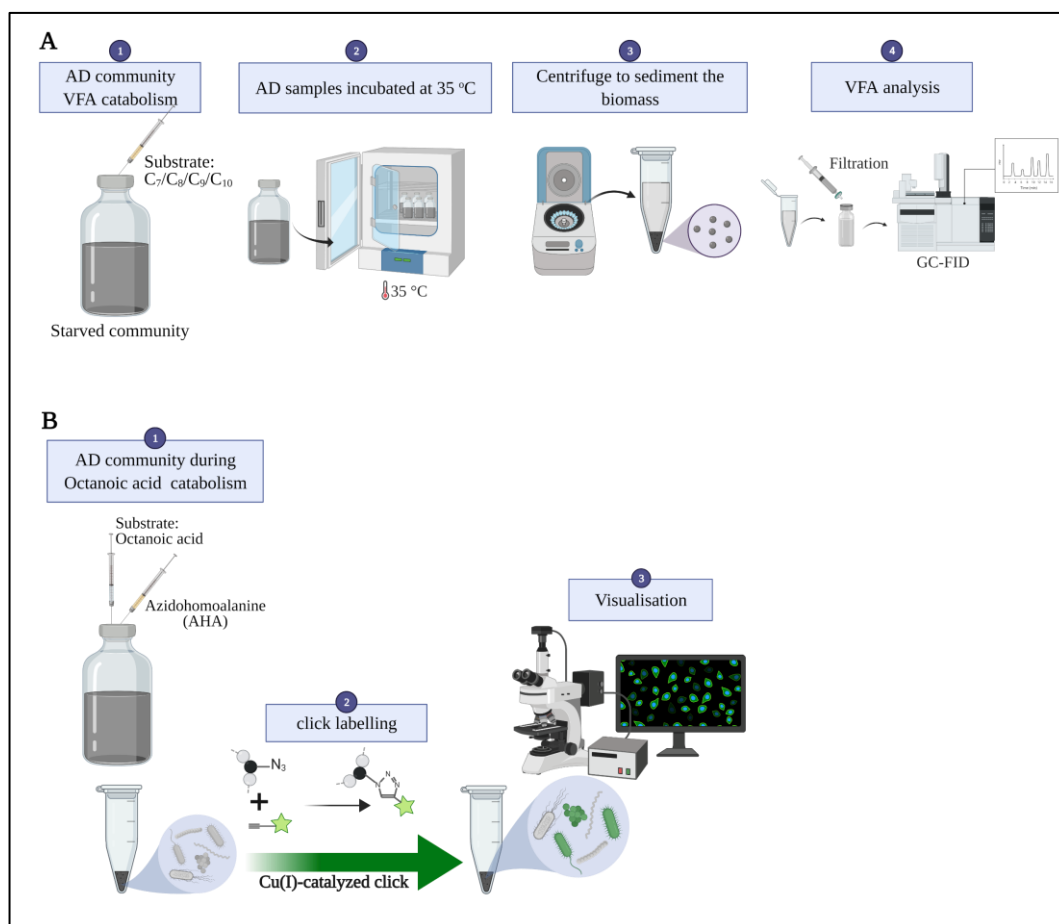
Figure 3.1 BONCAT workflow for monitoring and tracking active microbes in an AD community. Experimental design for (A) monitoring AD community medium-chain fatty acid catabolism dynamics via gas chromatography and (B) visualisation of metabolically active microbes in AD samples via BONCAT. Created using Biorender.com.

3.3 Results and Discussion

Starting material collected from AD systems contained an array of VFAs, such as acetic (C_2), propionic (C_3), *iso*butyric (iC_4), butyric (C_4), *iso*pentanoic (iC_5), pentanoic (C_5) and hexanoic (C_6) acids (Fig. 3.2-A). After incubation at 35 °C for 14 days of starvation, VFA content was undetectable by the GC-FID as shown in Figure 3.2. The VFA-starved community was used as the starting point for this study and allowed investigation of the community's response via VFA profile quantification and metabolically active cell labelling during medium-chain fatty

acid (MCFA) degradation. The dosage of MCFAs used in this study was proven to be tolerated by AD microbial communities (Koster and Cramer, 1987; Van Lier et al., 1993; Rinzema et al., 1994). Octanoic (C_8) and decanoic (C_{10}) acids were chosen as representative of MCFAs with an even number of carbons, while heptanoic (C_7) and nonanoic (C_9) acids were selected to represent an odd number of carbons MCFAs.
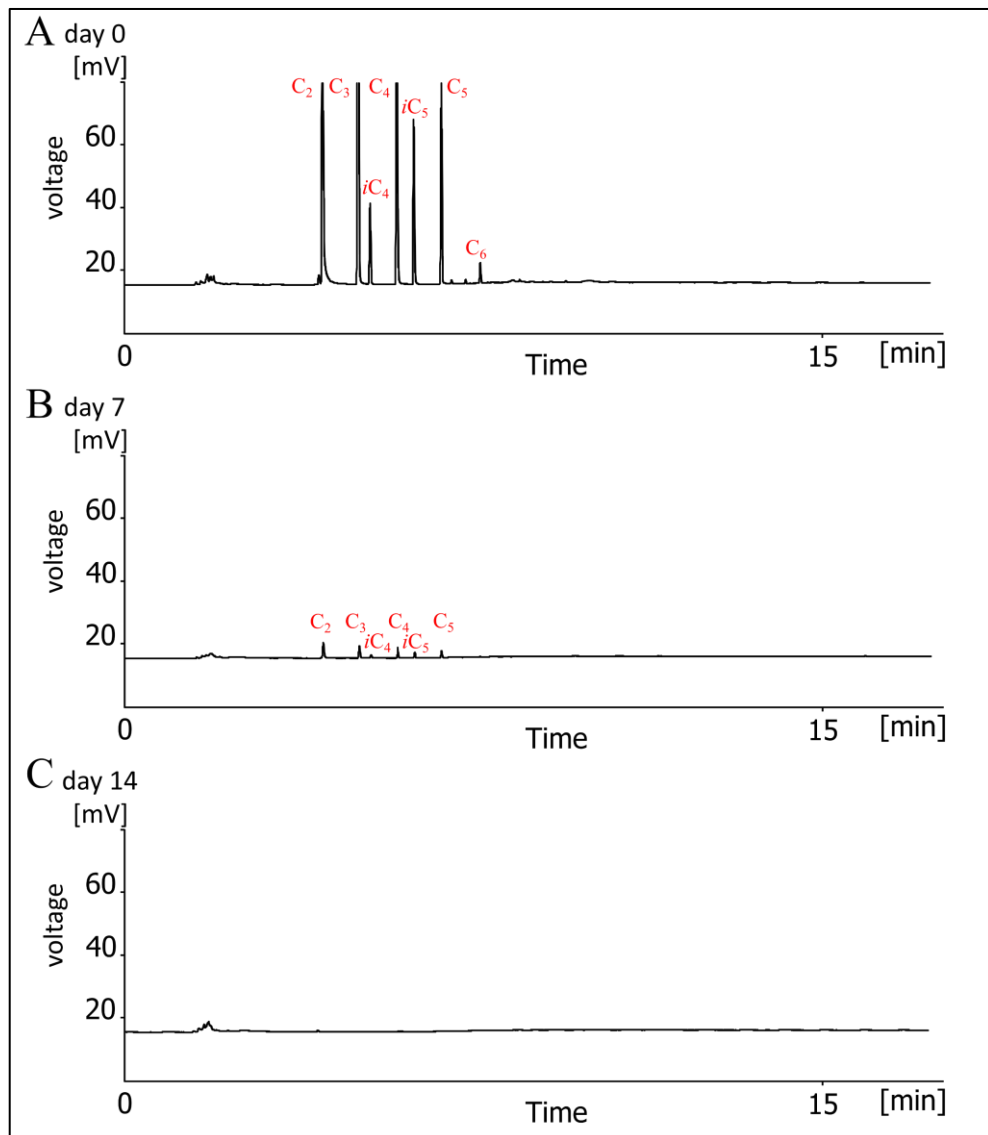
Figure 3.2 Volatile fatty acid (VFA) profiles during starvation of AD derived sludge. (A) fresh sludge; (B) after 7 days at 35 ℃; (C) after 14 days at 35 ℃. C_2-acetic; C_3-propanoic; iC_4-*iso*butyric; C_4-butyric; iC_5-*iso*pentanoic; C_6-hexanoic acid.

3.3.1 Catabolism of medium even-chain fatty acids by AD microbial community

Aliquots taken from the same batch of VFA-starved community were fed a single dose of ~10 mM octanoic (C_8) acid and incubated at 35 ^{o}C. VFA concentrations were quantified from samples taken every 6 h for 24 h post injection to establish the dynamics of octanoic acid degradation (Fig. 3.3 and 3.4). C_8 is transported exclusively by free diffusion into the cell as it has a permeability coefficient (0.173 cm/s) at least 100 times higher than water (Kamp and Hamilton, 2006). C_8 was then broken down in the cytosol. Hexanoic (C_6) and acetic (C_2) acid were observed in supernatants 6 h after addition of C_8 with a concomitant reduction in octanoic acid concentration. Neither octanoic nor hexanoic acids were detectable by 18 h, although acetic acid was still present. By 24 hours the VFA profile of the community had reverted to pre-addition levels, suggesting that octanoic acid was readily metabolised by this community. The major, persistent intermediates appear to be shorter, even-chain fatty acids (C_6, C_2), consistent with degradation through the β-oxidation pathway (Sousa et al., 2007b; Sousa et al., 2009). SCFAs ($C_4 - C_6$) presumably transported via porin channels (*e.g.* the porin OmpF) to cross the outer membrane and then diffuse across the cytoplasmic membrane in the non-ionized form (Clark and Cronan, 2005; Rodríguez-Moyá and Gonzalez, 2015) before being oxidised in the cytosol to C_2 acids.
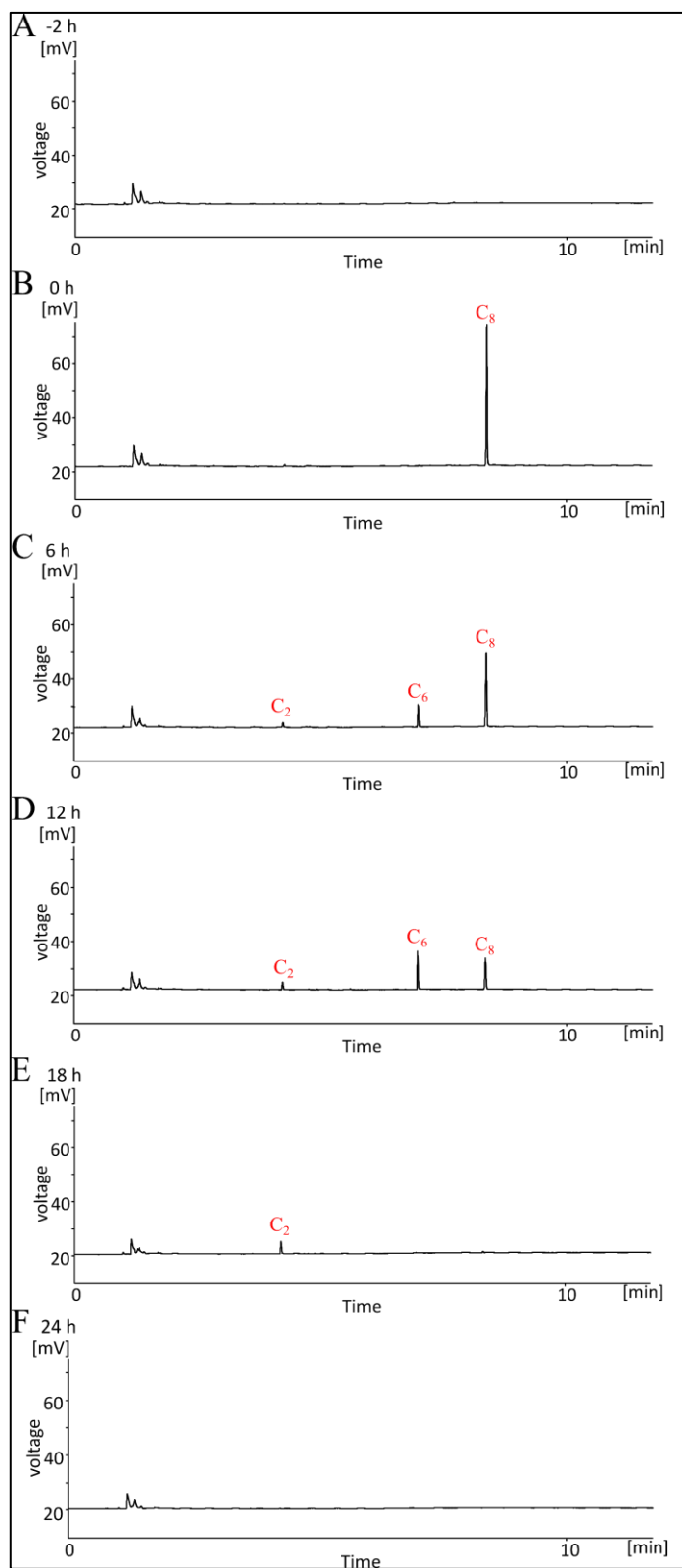
Figure 3.3 Comparison of volatile fatty acid (VFA) profiles during 24 hours of octanoic (C_8) acid catabolism by AD-derived sludge. (A) Starved sludge before C_8 spike; (B) after C_8 spike; (C-F) after 6 – 24 h at 35 ^{o}C. C_2-acetic; C_6-hexanoic; C_8-octanoic acid.
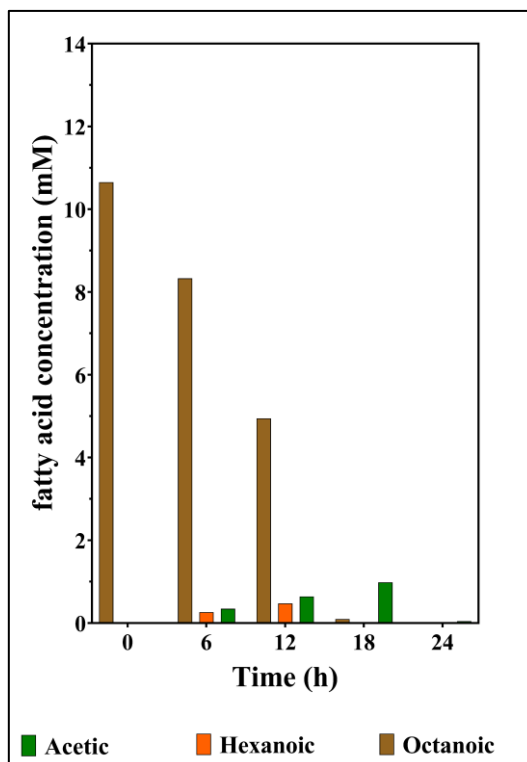
Figure 3.4. Time course of octanoic (C_8) acid degradation by AD-derived community.

Under the same conditions, 6 h after adding $\sim$ 12 mM decanoic (C_{10}) acid, shorter even-chain fatty acids (C_8, C_6, C_2) were detected as major intermediates with a corresponding reduction in C_{10} concentration (Fig. 3.5 and 3.6). By 24 hours, C_2 was augmented with simultaneous depletion in C_{10}, C_8 and C_6 concentrations relative to the 6 h timepoint. The VFA profile of the community had returned to pre-addition levels after 48 h. C_{10} was oxidised by the 'starved' AD-community through the sequential cleavage of 2-carbon acetate groups, consistent with the catabolism of C_8, via β-oxidation pathway that is catalysed by acyl-CoA dehydrogenase, enoyl-CoA hydratase, 3-hydroxyacyl-CoA dehydrogenase and β-ketothiolase (Fig. 1.8; Sousa et al., 2007b; Sousa et al., 2009). C_{10} (0.1394/h) was degraded much more slowly than C_8 (0.4400/h), consistent with Loehr and Roth (1968) that the degradation rate of fatty acids decreased with increased chain length as it needs to repeatedly re-enter β-oxidation pathway.
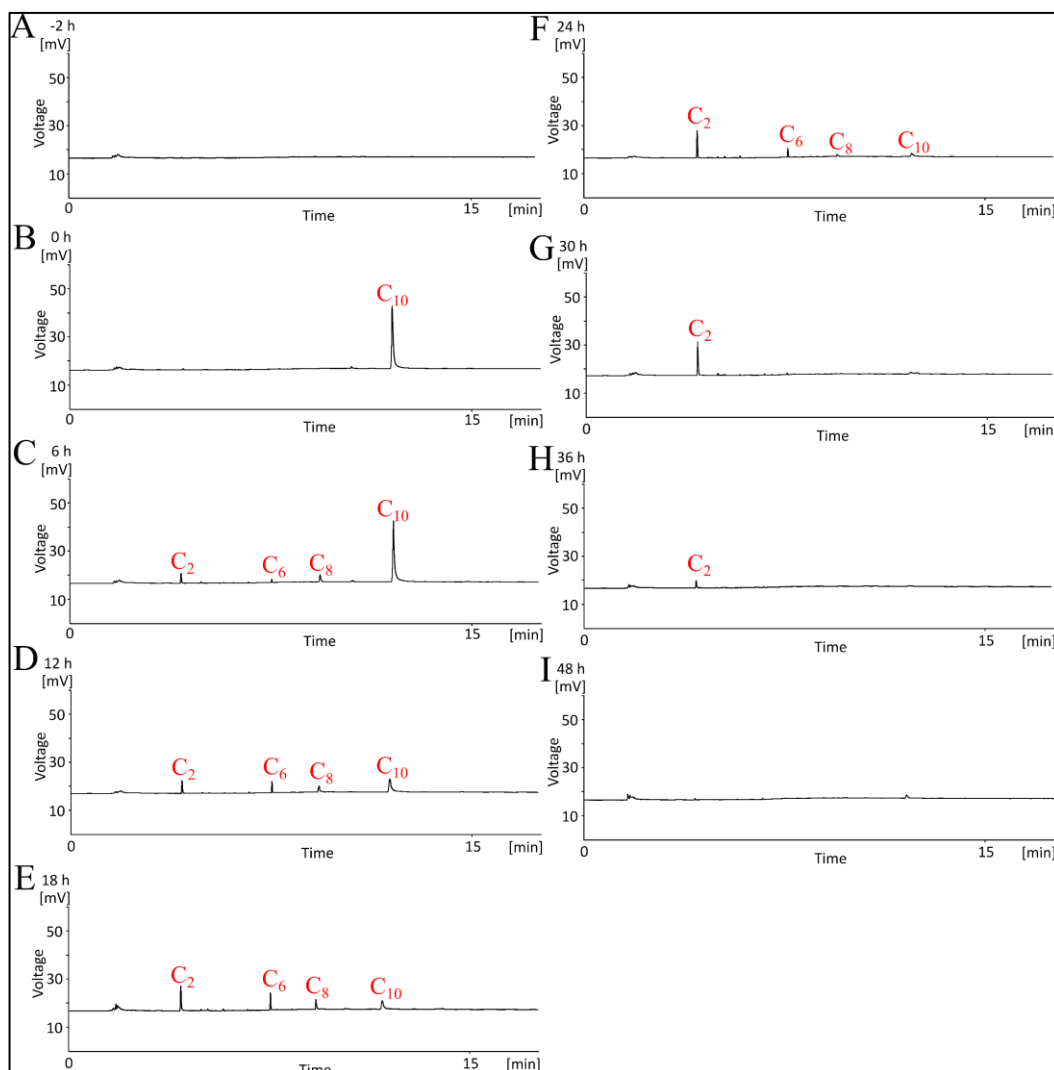
Figure 3.5 Comparison of volatile fatty acids (VFAs) profiles during 48 hours of decanoic (C_{10}) acid catabolism by AD-derived sludge. (A) Starved sludge before C_{10} spike; (B) after C_{10} spike; (C-I) after 6 – 48 h at 35 °C. C_2-acetic; C_6-hexanoic; C_8-octanoic acid; C_{10}-decanoic acid.
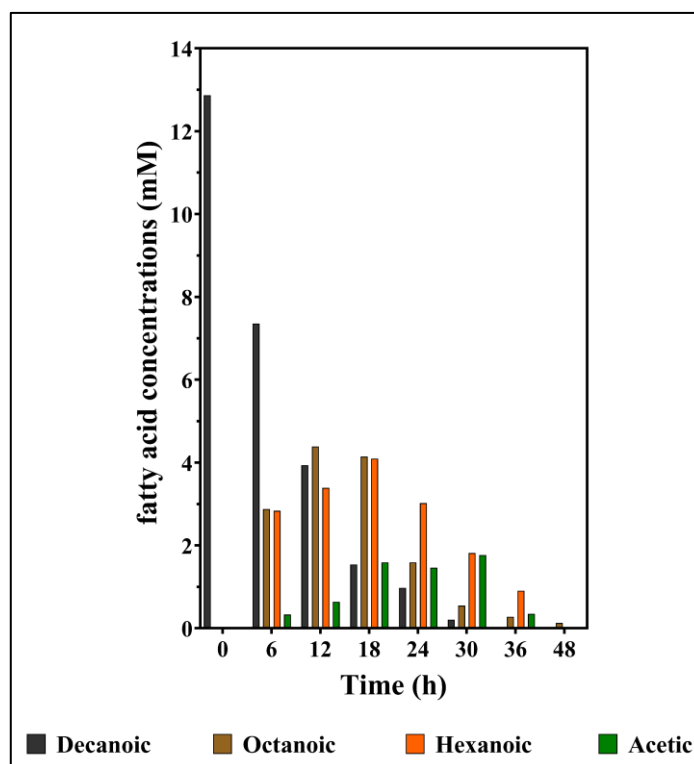
Figure 3.6 Time course of decanoic (C_{10}) acid degradation by AD-derived community.

Traces of butyric (C_4) acid were not observed during the catabolism of even-numbered MCFAs (C_8, C_{10}), contrasting with the Sousa et al. (2007b), seemingly due to higher degradation rate of C_4 compared to the major intermediates (C_2, C_6) detected in this study. Butyric acid degradation rate has been shown to be slowed down by the β-oxidation end-product, acetic acid when the concentration > 10 mM (Lin and Hu, 1993; Wang et al., 1999). Butyric acids were undetectable in the AD of C_8 and C_{10} acid, indicating that accumulated C_2 concentrations (max 0.98 mM and 1.76 mM in C_8 and C_{10} degradation, respectively) had no effect on C_4 oxidation rates by butyric acid degrading bacteria (*e.g.*, *Bacillus*, *Pseudomonas*, *Syntrophomonas* and *Syntrophus*; Fig. 3.7) which were expected to be abundant. Acetoclastic methanogens might have utilised the C_2 fraction produced during C_8 and C_{10} β-oxidation to generate methane (CH_4) alongside hydrogenotrophic methanogens that yield methane via CO_2 reduction using hydrogen atoms provided

from dehydrogenation in the course of β-oxidation. This syntrophic activity helps in keeping C_2 concentrations and H_2 partial pressure low enough for butyric acid degradation to take place (Wang et al., 1999).
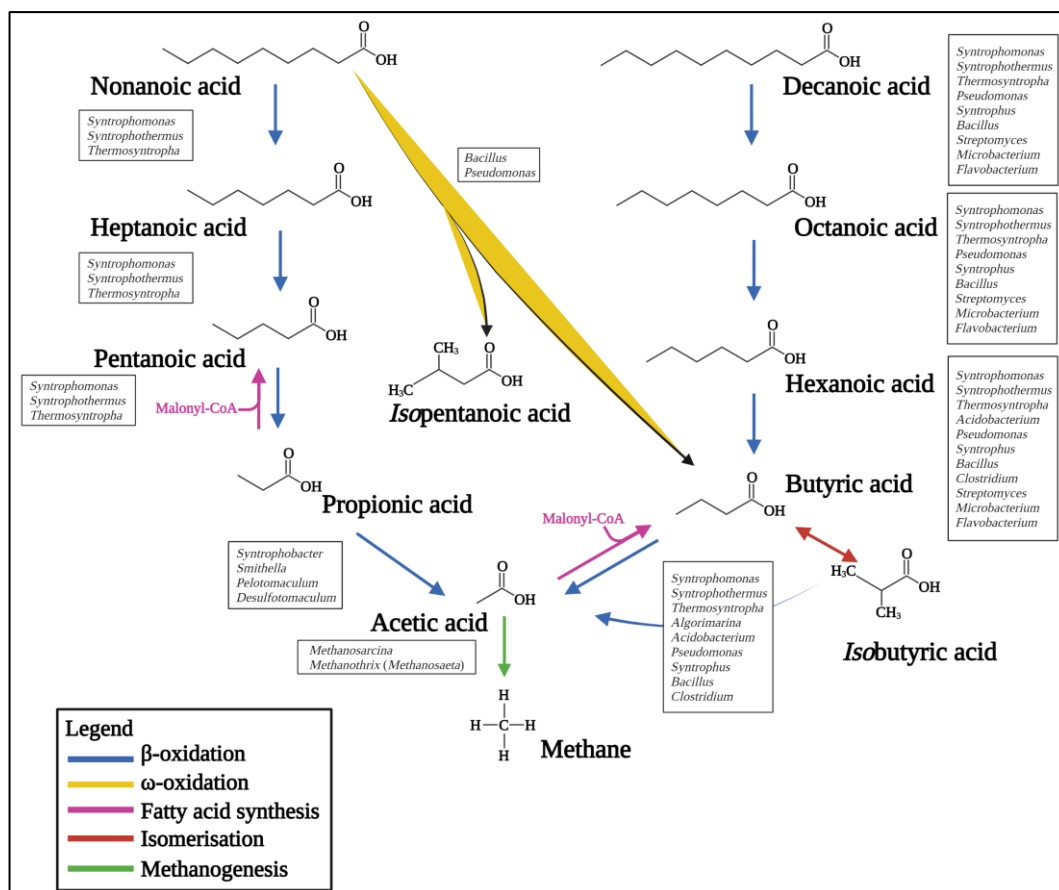
Figure 3.7 Various metabolic reactions and the predicted abundant genera in the AD of even- and odd-medium-chain fatty acids. Predicted abundant genera is based on the presence of genes encoding enzymes involved in fatty acid β-oxidation cycle inferred from their published genomes and KEGG pathway (Kanehisa et al., 2016). Created with Biorender.com.

Decanoic (C_{10}) acid is the first solid aliphatic carboxylic acid (melting point 31.5 $^{\circ}$C), unlike the first nine of the aliphatic carboxylic acids homologous series that are liquids at room temperature (Brondz, 2016). This feature makes it difficult to accurately add neat C_{10} into the sample. In this study, C_{10} was not dissolved in aqueous ethanol considering ethanol possible effects on the AD starved community

growth, viability, and metabolism. Therefore, to add neat C_{10} in liquid form, it was heated to 35 – 37 $^{\circ}$C and the transfer process was carried out in an incubator at 35 $^{\circ}$C. However, every time samples were collected and prepared for VFA analysis, crystalline C_{10} was observed in the liquid fraction as the processes occurred at room temperature. This cause variability in the C_{10} concentrations between replicates in this study, although the major intermediates profile appears to be consistent.

3.3.2 Catabolism of medium odd-chain fatty acids by AD microbial community

Aliquots of the starved AD-community were fed a single dose of ~ 20 mM of heptanoic (C_7) or nonanoic (C_9) acid to establish the dynamics of medium odd-chain fatty acids degradation. Accurately adding exact 50 mL volume of starved AD samples into 100 mL serum bottles was difficult as the samples contain solids and fine air bubbles. This makes it difficult to get exactly same final concentration of supplemented fatty acid, resulting in initial concentrations that are consistently higher than expected based on GC-FID measurement.

Heptanoic (C_7) acid was cleaved relatively slowly by the 'starved' AD-community (Fig. 3.8 and 3.9). Elevated concentrations of pentanoic (C_5), propionic (C_3) and acetic (C_2) acid were observed during the 96 h after addition with a concomitant reduction in C_7 concentration. This suggests that anaerobic microbial oxidation of C_7 mainly occurs by sequential cleavage of 2-carbon acetate groups via β-oxidation, resulting in C_2 and C_3 as final products. By 240 h, traces of C_2, C_3 and *iso*-forms (*iso*butyric (iC_4) and *iso*pentanoic (iC_5)) accumulated with simultaneous depletion of C_5 and C_7 concentrations. This indicates that C_3 and C_2

were formed faster by heptanoic- and/or pentanoic-oxidising bacteria (*e.g.,* *Syntrophomonas* and *Syntrophothermus*; Fig. 3.7) and degraded slower by propionic-oxidising bacteria (*e.g., Syntrophobacter*, *Smithella*, *Pelotomaculum* and *Desulfotomaculum*; Fig. 3.7) and acetoclastic methanogens (*e.g., Methanosarcina Methanothrix/Methanosaeta*; Fig. 3.7). The fact that C_2 was accumulated during this time shows that methanogens were much less active than the anaerobic bacterial populations. Consequently, it seemingly induces an increase in H_2 partial pressure. This could explain the strong accumulation of C_3, which its oxidation is highly sensitive to H_2 partial pressure (Gourdon and Vermande, 1987). Between 240 – 288 h, C_2 was rapidly degraded (decay rate 0.05/h) suggesting increased methanogen activity, while C_3 (9.05 mM), iC_4 (0.28 mM) and iC_5 (0.61 mM) were still present. This suggests that the increase in C_3 concentrations up to 9.05 mM is not inhibiting methanogens activity (Fig. 3.9). Subsequently, increased methanogenic activity could contribute to the decline in H_2 partial pressure, making C_3 degradation thermodynamically feasible (Fig. 3.8 J-M). Neither C_7 nor C_5 were detectable 288 h after addition. However, C_3 decomposition rate was slow as it needs 204 h to be fully metabolised by the AD community. Oxidation of C_3 is known to require three further enzymes (propionyl-CoA carboxylase, methylmalonyl-CoA epimerase and methylmalonyl-CoA mutase) in addition to those required for even-chain fatty acid degradation. After 492 hours, the community VFA profile had reverted to pre-addition levels which indicates C_7 was fully degraded by this community.

Interestingly, traces of some *iso*-forms, *e.g. iso*-butyric (*i*C4) and *iso*-pentanoic (*i*C5), were observed after 240 h of the anaerobic degradation of C_7 (Fig. 3.8 and

3.9). iC4 and iC5 were found in low concentrations, 0.31 and 0.89 mM, respectively (Fig. 3.9). The starved-AD community VFA profile indicates that iC4 and iC5 were products of microbial activity during C_7 degradation, although other reports have suggested that these short-chain methyl-branched VFAs could be generated by microbial (*e.g., Bacteroides*, *Corynebacterium*, *Megasphaera, Propionibacterium* and *Staphylococcus* species) catabolism of branched-chain amino acid valine and leucine/*iso*leucine, respectively (Allison, 1978; Thierry et al., 2004; James et al., 2013). However, given the evidence here (Fig. 3.9) showing increased branched-chain VFAs formation by AD-starved communities during C_7 degradation, it is worth noting that this phenomenon occurred when C_2 (7.14 mM) and C_3 (8.07 mM) concentrations were high. This could induce microorganisms, such as facultatively anaerobic *Salmonella*, to start *de novo* fatty acids synthesis by using either acetyl-CoA (converted from acetic acid) or propionyl-CoA (converted from propionic acid) as a starter unit for condensing with malonyl-CoA, resulting in 4- or 5-carbon SCFAs (Park et al., 2020).
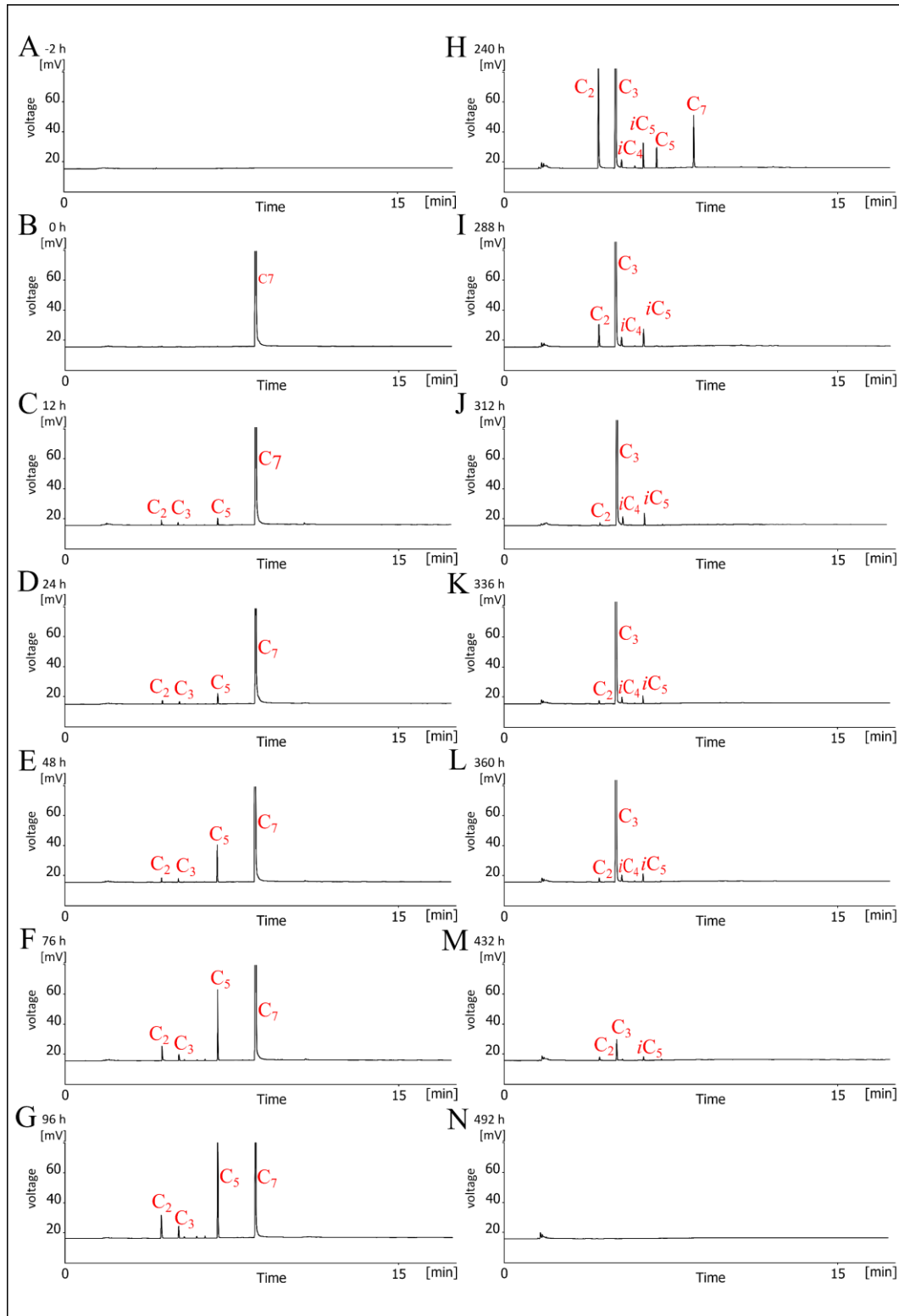
Figure 3.8 Comparison of volatile fatty acids (VFAs) profiles during heptanoic (C_7) acid catabolism by AD-derived sludge. (A) Starved sludge before C_7 spike; (B) after C_7 spike; (C-I) after 12 – 492 h at 35 ^{o}C. C_2-acetic; C_3-propioic; iC_4-*iso*butyric; iC_5-*iso*pentanoic; C_5-pentanoic; C_7-heptanoic acid.
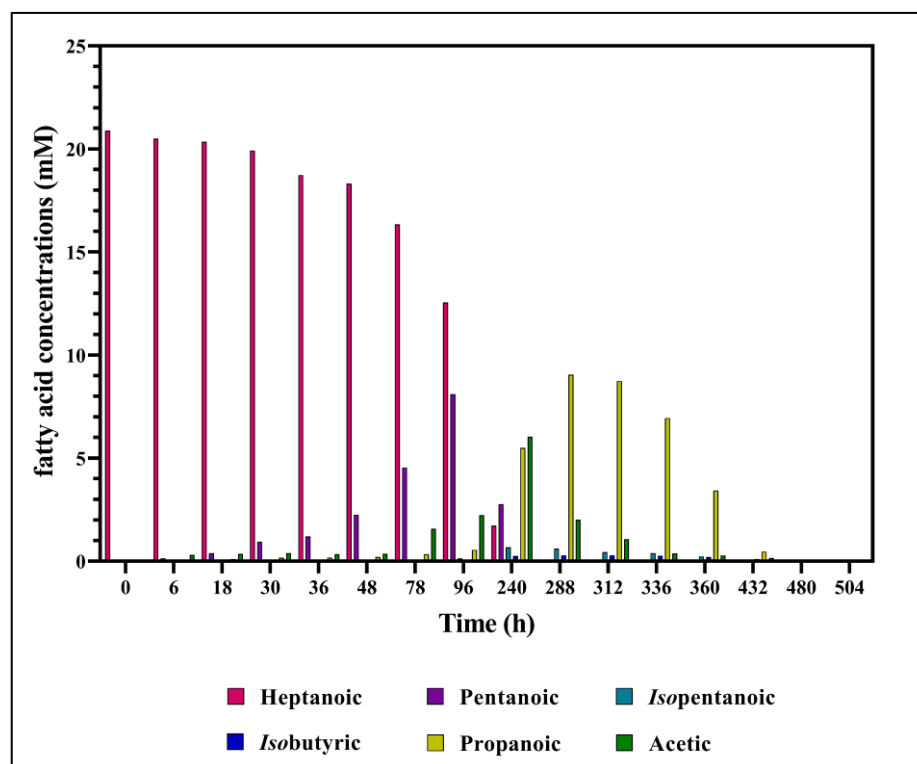
Figure 3.9 Time course of heptanoic (C_7) acid degradation by AD-derived community.

The catabolism of normal-form VFAs (C_5; decay rate 0.03/h) was faster than their respective *iso*-forms (iC_5; decay rate 0.02/h) consistent with Wang et al. (1999) as the trace of iC_5 was still observed until 336 h after addition. The anaerobic degradation of C_4 and iC_4 presumably goes through reciprocal isomerisation (Matthies and Schink, 1992) before being catabolised via β-oxidation to yield C_2 acids. However, there is no reciprocal isomerisation between C_5 and iC_5 (Wang et al., 1999) prior to β-oxidation to produce C_3 and C_2 acids. In the case of iC_5, it would also have generated C_2 and acetone (CH_3COCH_3) or isopropyl alcohol ($CH_3CH(OH)CH_3$) if it is catabolised via β-oxidation. But only C_2 was observed (Fig. 3.8 and 3.9). The GC-FID method used in this study was not capable of identifying acetone and isopropyl alcohol as both have lower boiling point (56 °C and 82.5 °C, respectively) than the temperature gradient applied for sample elution (see sec. 2.3.1). Thus, it is not possible to fully elucidate the degradation route of

101

iC$_5$. Acetone and isopropyl alcohol still could be tested using GC-FID by applying different temperature gradient, for example the oven temperature is programmed to 40°C (for 2 min), followed by an increase of 5°C/min until 200°C (Pontes et al., 2009).

Nonanoic (C$_9$) acid was more slowly degraded (decay rate 0.0121/h) to C$_3$ and C$_2$ compared to C$_7$ caused by its longer carbon chain. Neither C$_7$ or C$_5$ acids were detectable by 96 h after addition of C$_9$, although C$_3$ and C$_2$ as final products and the major persistent intermediate metabolites iC$_5$, C$_4$ and iC$_4$ acids were present (Fig. 3.10). It was found that microbial β-oxidation activity was low when C$_9$ > 10 mM (Fig. 3.11). The condition seemingly induced active microorganisms to use an alternative oxidation pathway, ω-oxidation, which normally occurs for medium chain fatty acids catabolism (Schönfeld and Wojtczak, 2016). Some Bacteria, e.g. *Bacillus megaterium* and *Pseudomonas oleovolans*, have been reported to be able to oxidise free fatty acids at the ω, ω-1, ω-2, and ω-3 positions in the presence of NADPH and O$_2$ (Fig. 3.7; Miura and Fulco, 1975) which forms dicarboxylic acid and/or oxo-compounds. The formed dicarboxylic acids undergo β-oxidation usually from both ends of the fatty acid chain (bilateral β-oxidation) to shorter dicarboxylic acids, while the oxo-compounds are cleaved into two compounds via thiolysis (Miura, 2013). This explains the presence of low concentrations of 4- and 5-carbon SCFAs during this period, which is consistent with expected low levels of oxygen in the system likely introduced while transferring aliquots into the 100 mL serum bottles. The very active ω-oxidation consumes any remaining oxygen inside the system, resulting in an oxygen-free system, where anaerobes optimally thrive. The appearance of iC$_4$ and C$_4$ is consistent with Wu et al. (1994) and Wang et al. (1999)

as anaerobic degradation of C_4 and iC_4 presumably goes through reciprocal isomerisation (Matthies and Schink, 1992) before being catabolised via β-oxidation to yield C_2 acids. Between 264 – 504 h after addition, traces of C_7 and C_5 were observed as a result of sequential cleavage of 2-carbon acetate groups from C_9 and produced C_3 and C_2 as final products. C_9 concentration was 3.22 mM and 0.55 mM at 264 and 504 h after addition, respectively (Fig. 3.11). This indicates that microbial β-oxidation of C_9 is more favourable when C_9 concentration is < 5 mM. No C_9, C_7 and C_5 were detected at 1032 h after addition, although C_2, C_3, iC_5 and iC_4 were still present. The catabolism of normal-form VFAs (C_4 and C_5) were faster than their respective *iso*-forms consistent with Wang et al. (1999) and C_7 experiment as the traces of both *iso*-forms were still observed until 1080 h after addition (Fig. 3.10 L-M). C_3, one of the β-oxidation OCFA products, was catabolised slowly by the community. Presumably, C_3 needs to undergo a unique metabolism so that it is sequentially carboxylated, isomerised and rearranged for it to be converted to succinic acid $((CH_2)_2(CO_2H)_2)$ (Galivan and Allen, 1968), which is then further oxidised via oxaloacetate and pyruvate to yield acetic acid (Schink, 1985). Acetic (C_2) acid then cleaved by the acetotrophic methanogens to release methane and carbon dioxide. The VFA profile of the community had returned to pre-addition levels after 1280 h. The GC-FID method used in this study was less sensitive in detecting succinic acid (less volatile with boiling point 235 °C) than it is for VFAs because succinic acid contains fewer $-CH_2-$ and $-CH_3$ groups than propionic and butyric acid (Playne, 1985). Alternatively, raising the maximum column temperature to 250 °C may improve the detection of this relatively high boiling point compound.
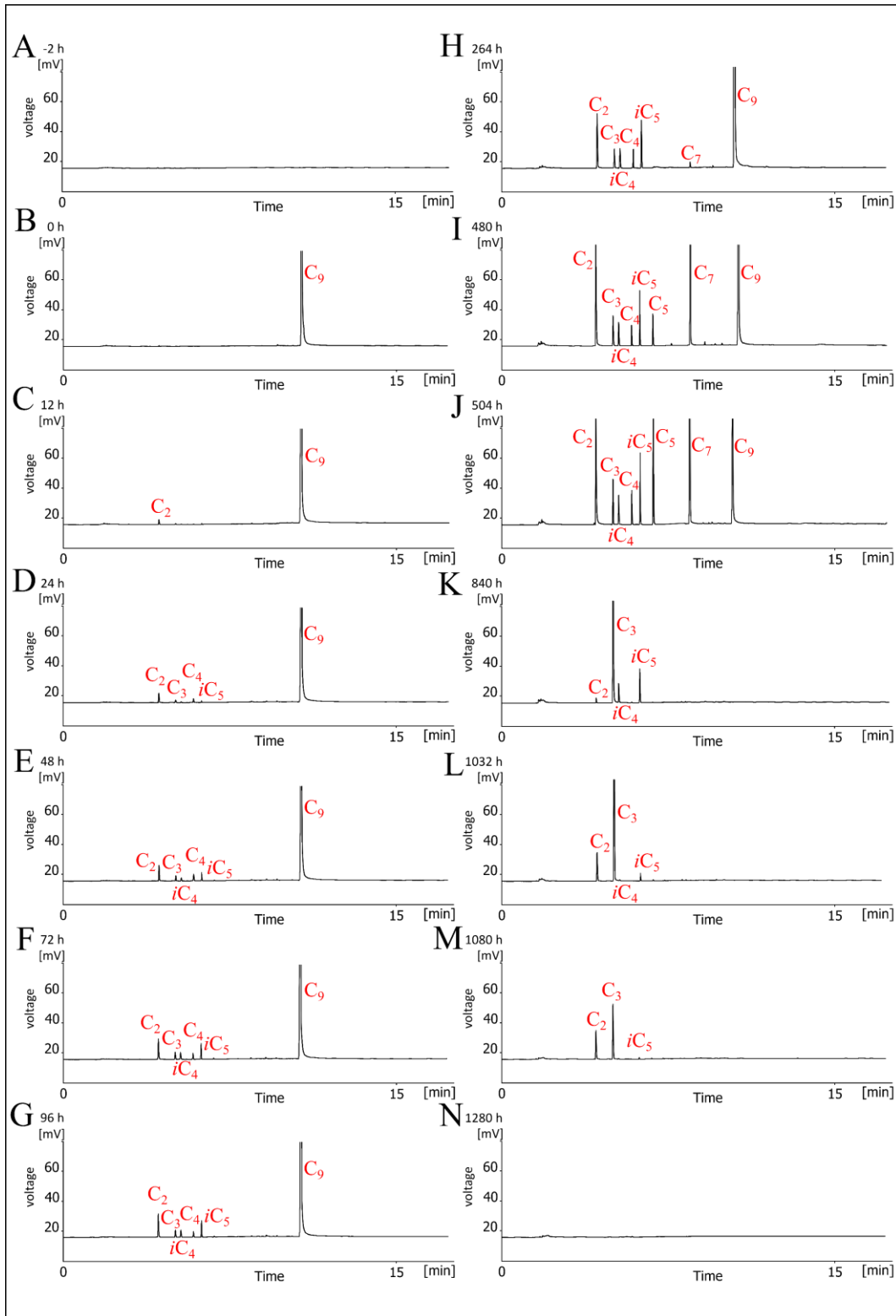
Figure 3.10 Comparison of volatile fatty acid (VFA) profiles during nonanoic (C_9) acid catabolism by AD-derived sludge. C2-acetic; C3-propionic; iC_4-*iso*butyric; C_4-butyric; iC_5-*iso*pentanoic; C_5-pentanoic; C_7-heptanoic; C_9-nonanoic acid.
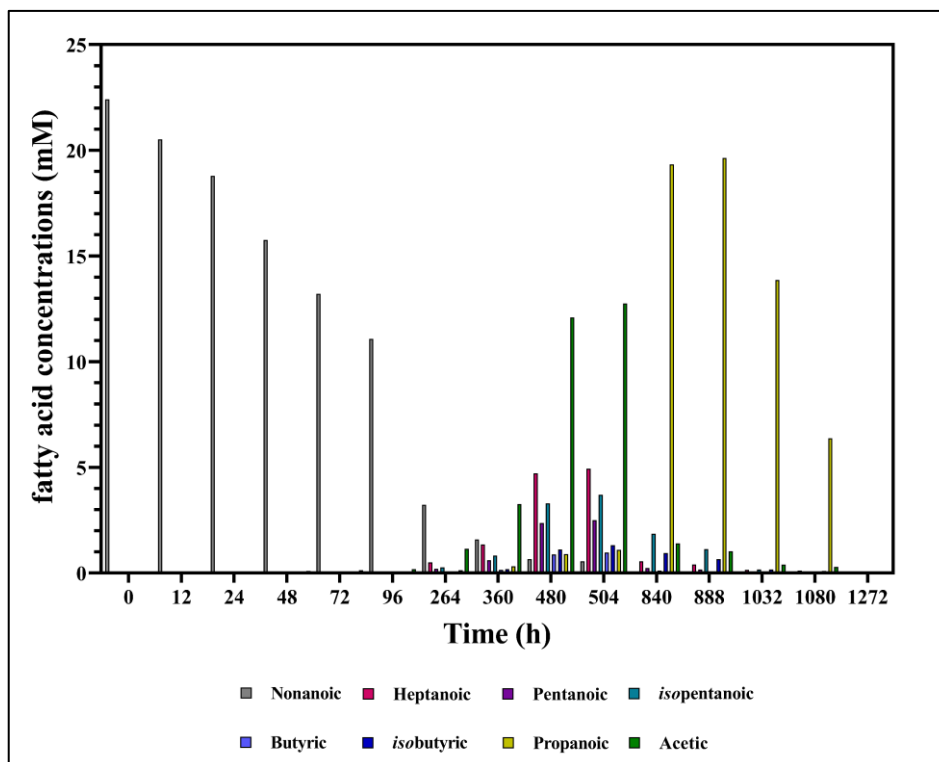
Figure 3.11 Time course of nonanoic (C_9) acid degradation by AD-derived community.

Overall, this study shows that anaerobic degradation of odd-chain fatty acids (C_7 and C_9) were so much slower than the oxidation of even-chain fatty acids (C_8 and C_{10}). The slow anaerobic degradation rate of OCFAs is strongly linked to the initial concentration of the supplemented MCFA, OCFA degraders abundance and the presence of propionic (C_3) acid. The slow rate of C_7 and C_9 anaerobic oxidation during initial periods indicates that OCFA oxidisers are present in very low abundance after the microbial populations underwent starvation. Moreover, the initial concentrations of C_7 and C_9 were double the initial concentrations of ECFAs applied in this study. Both starvation and overloading may have a strong effect on AD microbial populations, especially the sensitive low abundance syntrophic bacteria and methanogens, which potentially leaves a relatively low abundance of robust bacteria/methanogens as starters. Microbes (bacteria/methanogens) with a

105

higher tolerance of stress may outcompete less tolerant microbes and display slow anaerobic oxidation of OCFAs initially. This may then lead to an enhanced environmental condition for syntrophic bacteria to became more abundant and be accompanied by a boost of acetoclastic and hydrogenotrophic methanogen activity. Furthermore, the biodegradation of C_3 was found to be the slowest compared to other intermediate compounds and limit the anaerobic degradation rate of OCFAs by the VFA-starved AD community. Propionic acid degraders are syntrophs which cohabit with autotrophic methanogens to overcome the thermodynamic barrier created by high H_2 partial pressure. Currently, only 10 species of propionic acid degraders have been identified, e.g. *Syntrophobacter wolinii*, *S. pfennigii*, *S. fumaroxidans*, *S. sulfatireducens*, *Smithella propionica*, *Pelotomaculum schinkii*, *P. thermoproopionicum*, *P. propionicum*, *Desulfotomaculum thermocisternum* and *D. thermobenzoicum* subsp. *Thermosyntrophicum*. Seemingly, these propionic degraders grow relatively slowly due to their fastidious metabolism via methyl-malonyl-CoA (yields acetyl-CoA, NADH and $FADH_2$) and being outcompeted by other bacteria in the presence of alternative fatty acids intermediates in this study.

3.3.3 Obtaining a temporal snapshot of the active cell fraction in AD system using BONCAT

The presence of intermediates indicates that either the fatty acid is able to enter β-oxidation faster than the intermediates can be utilised, saturating microorganism capability to metabolise it, or that different bacteria are required to oxidise various intermediates, causing cross-feeding in the AD system. This kind of microbe-microbe interaction could influence the fatty acid metabolism carried out within the AD community. To decipher the diversity, role and function of AD microbial

communities and link these to their in-situ fatty acid degradation activity, techniques that distinguish active microbes from extracellular DNA and dormant cells is required. Here, BONCAT was used to obtain a temporal snapshot of the active cell fraction during fatty acid catabolism.

3.3.3.1 Determine the minimum incubation time for fluorescence detection of BONCAT signals in AD sample

Initially, AHA-labelling incubations with AD samples were performed to establish a protocol for the fluorescence labelling of AHA-containing proteins in chemically fixed cells. A labelling protocol (sec. 2.4.1 – 2.4.2) was adjusted in accordance with Hatzenpichler and Orphan (2015). AHA-containing *Escherichia coli* were used to test Cu(I)-catalyzed click reactions with Alexa-488 alkyne fluorescent dyes (Fig. 3.12) and showed successful fluorescent labelling with a slide-immobilised biomass approach.

Time course experiments with AD samples were performed to determine the minimum time required for detection of AHA-labelled cells by BONCAT. VFA-starved AD communities were incubated in the presence or absence of 1 mM AHA for 30, 60, 120 and 240 minutes (Fig. 3.13). These incubation times were tested based on the predicted growth rate of prokaryotes in an anaerobic system, *e.g.* the human gut, with minimum doubling times ranging from 0.5 – 5 h for fast-growing phyla (*e.g.* Actinobacteria, Bacteroidetes, Firmicutes and Proteobacteria) and > 5 h for slow-growing phyla (*e.g.* Chloroflexi, Cyanobacteria and Planctomycetes) (Gibson et al., 2018; Weissman et al., 2021). Thus, all incubations were performed ≤ 1 predicted generation time of the fast- and slow-growing microbes in the system.

Intact cells were fixed in paraformaldehyde (PFA), fluorescently labelled with Alexa-488 alkyne fluorescent dyes via click chemistry and DAPI-stained after incubations.
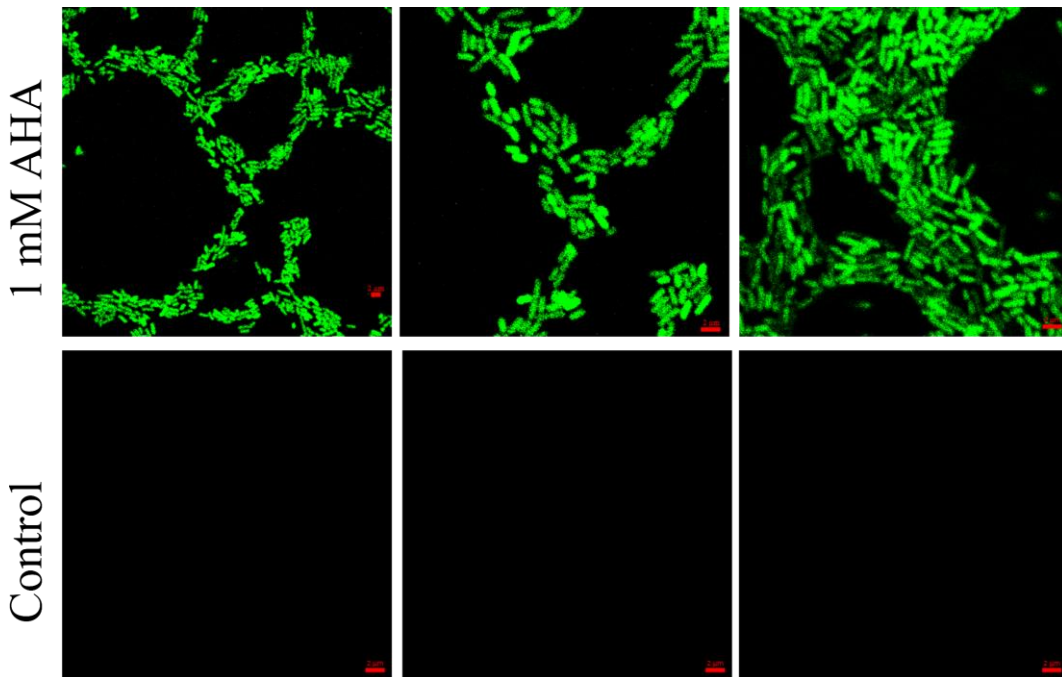
Figure 3.12 BONCAT of translationally active *E. coli* cells. All scale bars applied are equal to 2 μm. BONCAT signals (green) were recorded after *E. coli* incubation in the presence or absence of 1 mM AHA for 15 minutes at 35 °C. BONCAT signals were taken at identical exposures time using confocal microscope LSM 710 (Axio Imager 2, ZEISS) with objective Plan-Apochromat 63x/1.4 Oil DIC M27.
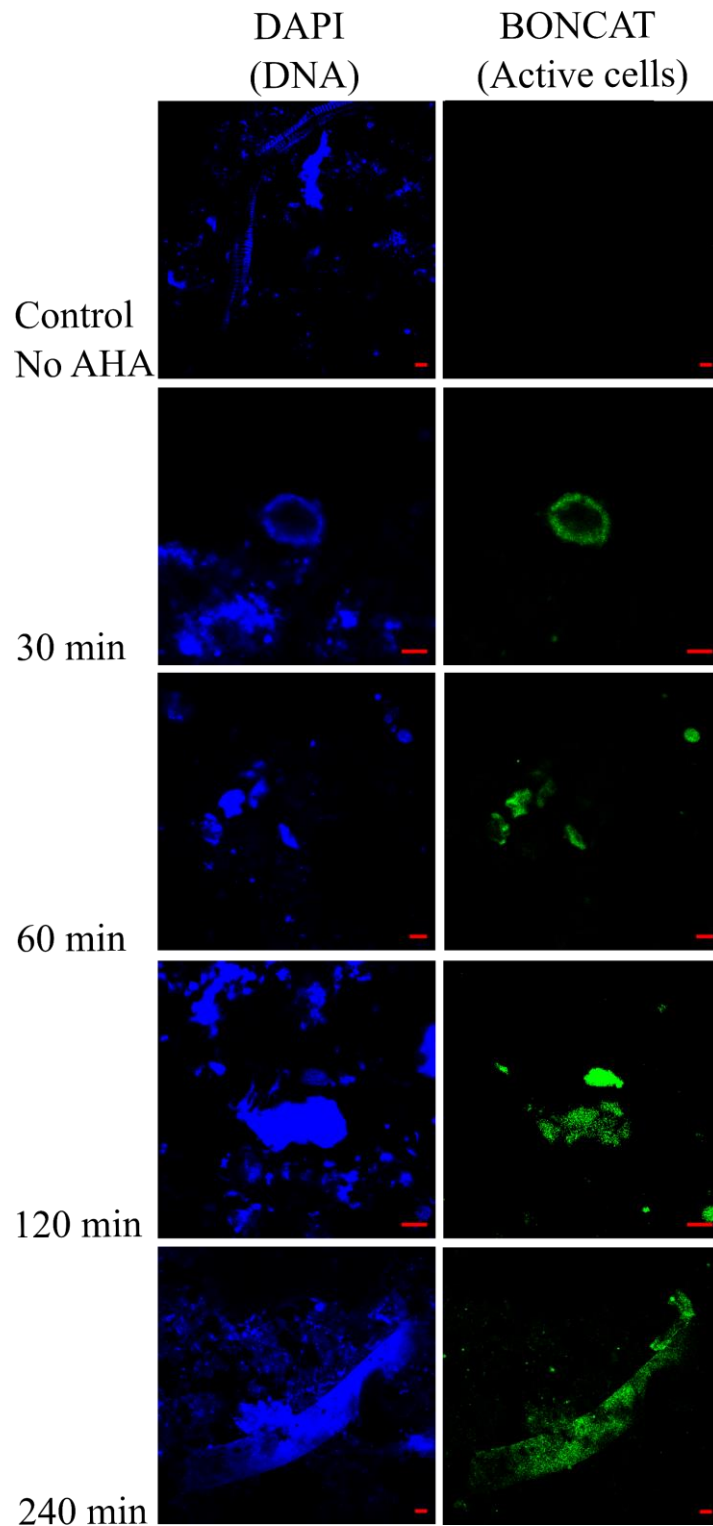
Figure 3.13 Fluorescence labelling of AD microbial community over time. All scale bars applied are equal to 5 μm. BONCAT signals (green) were recorded after AD microbial community incubation in the presence or absence of 1 mM AHA at 35 °C. BONCAT signals were taken at identical exposures time using confocal microscope LSM 710 (Axio Imager 2, ZEISS) with objective Plan-Apochromat 63x/1.4 Oil DIC M27.
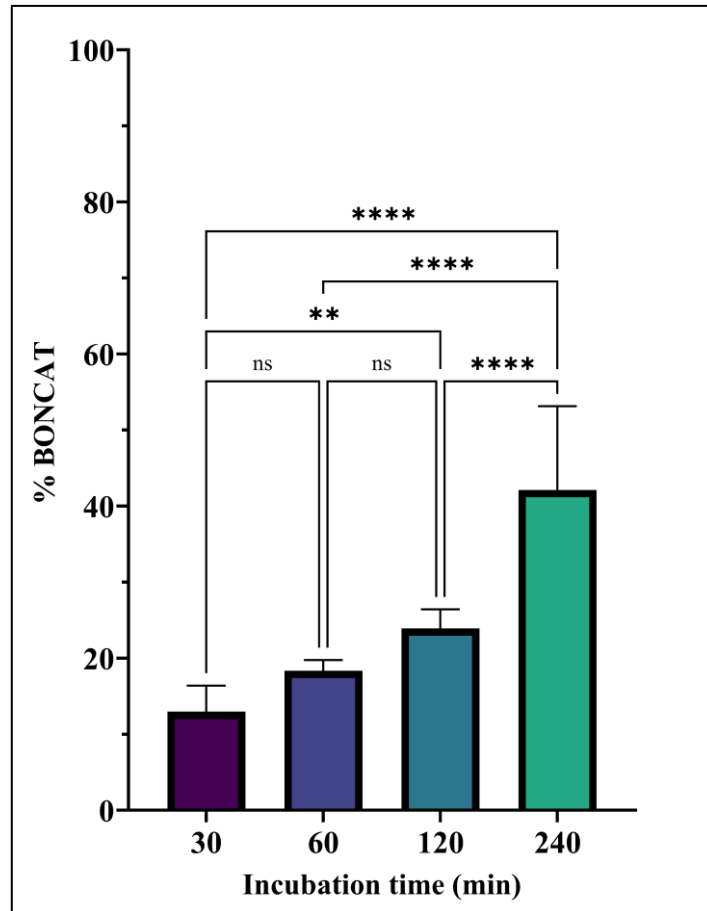
Figure 3.14 The percentages of BONCAT positive cells out of DAPI positive cells in AD-derived sludge over time in the presence of 1 mM AHA (n =10). The significance of differences was analysed by Tukey test (**** P <0.001; ns, not significant) and performed using GraphPad Prism software version 9.0.2.

Qualitative and quantitative analysis based on the green fluorescence intensity revealed that the fluorescence signal as a result of AHA incorporation by AD communities increased over time relative to DAPI fluorescence signals in all samples (Fig. 3.13 – 3.14). Flocs of microbial cells were observed to be metabolically/translationally active during the incubation period, which was verified by the no AHA control (Fig. 3.13), suggesting BONCAT applicability on AD samples. 1 mM of AHA applied resulted in ~ 2.5 µM bioavailable amino acid in the sample due to its ~ 0.25% MetRS activation rate compared to Methionine (Kiick et al., 2002; Hatzenpichler et al., 2014), which appears to be sufficient to

induce uptake by active anaerobic microbes. After 30 minutes incubation, the AD community exhibited ± 13% BONCAT-fluorescence signal, which increased significantly to ± 42% after 240 minutes of incubation. The BONCAT labelled AD community increased after 60 and 120 minutes of incubation, however the difference was not significant (Fig. 3.14).

In this study, 30 minutes of incubation with 1 mM AHA was chosen as the minimum incubation time because it is sufficient to label > 10% of AD microorganisms while minimising excessive substitution of L-Methionine with AHA, which could impede cellular machinery efficiency and reduce the risk of system disruption (Hatzenpichler et al., 2014). Furthermore, a short incubation time reduces the possibility of overlapping labelling, enhancing the precision of this technique in monitoring and identifying active microorganisms in samples at a certain interval. Unfortunately, estimation of labelling efficiency was not feasible due to the nature of AD samples, which were compact and contained densely packed cells which were not trivial to separate from the other materials in the samples. The inability to label the majority of the cells in AD samples was presumably due to the nature of slow-growing anaerobic microbes, which could not synthesise enough nascent AHA-containing proteins to be detected by the fluorescence microscope. Furthermore, because the most crucial limitation of BONCAT is its reliance on uptake mechanisms, it cannot be ruled out that some anaerobic microbes seemingly did not take up AHA into their cells or proteins due to the lack of appropriate transporters, the high selectivity of their MetRS, or a high Met/AHA ration in the cytoplasm (Hatzenpichler et al., 2014).

3.3.3.2 Tracking active microbial fraction in AD samples via BONCAT

To track the active microbes during specific substrate degradation by the 'starved' community, it was fed a single dose of ~10 mM octanoic acid and incubated at 35 $^{\circ}$C for 24 h. Octanoic acid was chosen because it was degraded faster by the VFA-starved AD community and has reproducible results compared to other MCFAs substrate used in this study. Subsampling for VFA analysis and click chemistry-mediated fluorescence labelling was conducted every 6 h following the consistent VFA degradation profile as shown in Fig. 3.3 and 3.15.

The VFA profile shows that octanoic acid (C_8) was metabolised by the 'starved' community through the sequential cleavage of two-carbon fragments. Consequently, the major intermediate metabolites of octanoic degradation are other shorter chain fatty acids, like hexanoic (C_6) and acetic acid (C_2) (Fig. 3.15). This cascade degradation affirms that the breakdown of even-chain fatty acids by the 'starved' inocula in mesophilic conditions generates acetate and indicates the occurrence of β-oxidation.
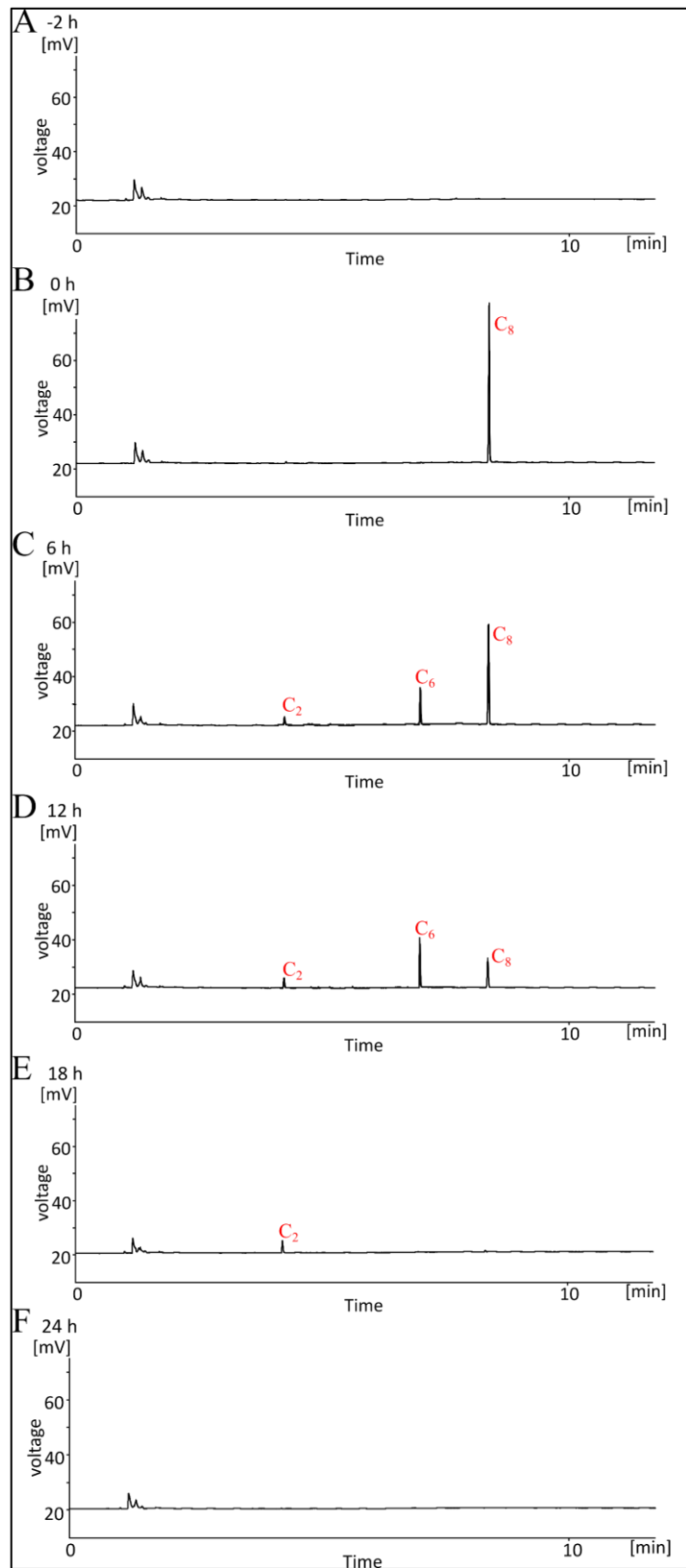
Figure 3.15 Comparison of volatile fatty acids (VFAs) profiles during octanoic (C_8) acid catabolism by VFA-starved AD community. (A) Starved sludge before C_8 spike; (B) after C_8 spike; (C-F) after 6 – 24 h at 35 °C. C_2-acetic; C_6-hexanoic; C_8-octanoic acid.

Visualisation of the active cells during octanoic degradation was done following an established BONCAT protocol with slide-immobilized biomass (Hatzenpichler and Orphan, 2015; sec. 2.4.1 – 2.4.2). Any microorganisms that are actively synthesising new proteins during C_8 amendment and degradation will incorporate AHA into their nascent proteins. The active cells were visualised using fluorescence microscopy by attaching Alexa-488 alkyne via Cu(I)-catalysed click chemistry to the nascent proteins which have incorporated the synthetic amino acid. DAPI-stained clusters that show high intensity 488 nm fluorescence are referred to as BONCAT-active microbes (Fig. 3.16). The VFA-starved AD community exhibited $\pm$ 15% BONCAT-fluorescence signal before C_8 addition (0 hour; Fig. 3.17), suggesting the presence of metabolically active anaerobic microbes (bacteria/archaea) that are able to survive after two weeks of starvation. Apparently, these surviving cells had to scavenge nutrients from what was in the environment, *e.g.* from cellular detritus. When nutrients were available again after C_8 addition, survivor microbes were forced to adapt to utilise C_8 as a source of energy, despite the fact that it is not normally present in the system (Fig 3.2-A). These persistent bacteria/archaea populations may play a role in promoting and facilitating C_8 anaerobic oxidation, resulting in an improved environment for syntrophic bacteria and methanogenic archaea to become more abundant. The proportion of cells stained by BONCAT increased significantly 12 – 18 h after addition and decreased when C_8 was entirely metabolised by the AD community (Fig. 3.17), which confirm the existence of SCFA intermediates during octanoic acid oxidation (Fig. 3.15). This observation demonstrates the capability of BONCAT to visualise and monitor the fraction of *in situ* active microbes in an AD system.

| | DAPI
(DNA) | BONCAT
(Active cells) |
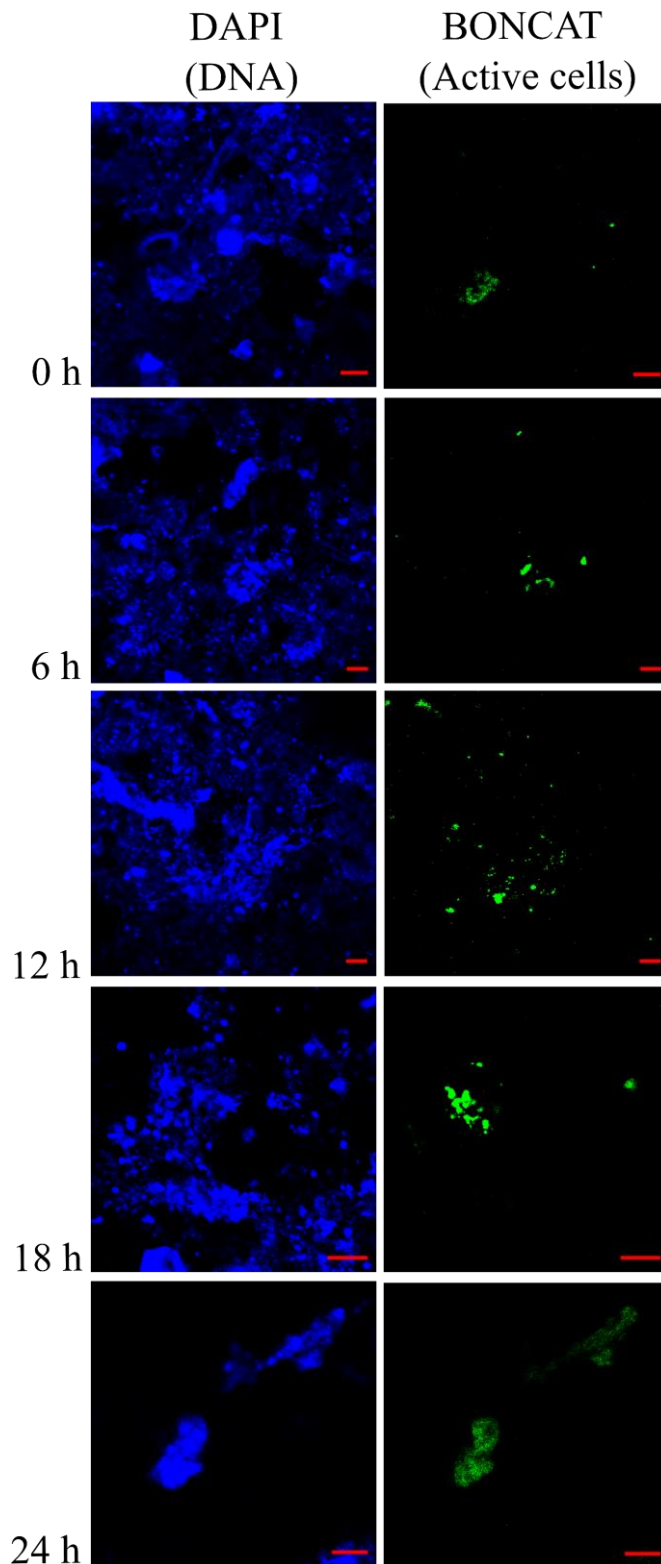
Figure 3.16 Visualisation of BONCAT-isolated active microbes in AD-derived sludge during octanoic acid degradation. Scale bars indicate 10 µm in all images. BONCAT signals were taken at identical exposures time using confocal microscope LSM 710 (Axio Imager 2, ZEISS) with objective Plan-Apochromat 63x/1.4 Oil DIC M27.

115
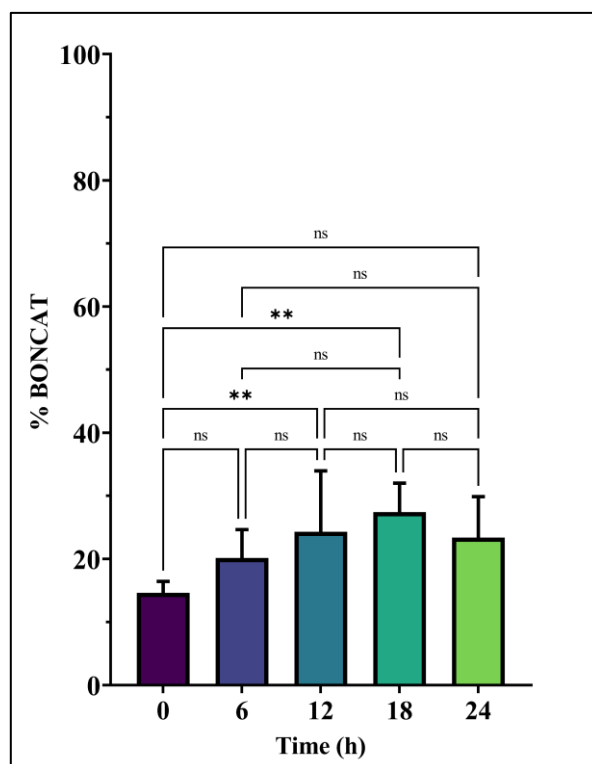
Figure 3.17 The percentages of BONCAT positive cells out of DAPI positive cells in AD-derived sludge during octanoic acid degradation at 35 $^{\circ}$C (n =10). BONCAT was performed for 30 minutes in the presence of 1 mM AHA. The Tukey test was used to determine the significance of differences (** P = 0.002; ns, not significant) and performed using GraphPad Prism software version 9.0.2.

3.4 Conclusion

This study demonstrated the dynamics of ECFAs and OCFAs catabolism by VFA-starved AD communities via the β-oxidation pathway. ECFAs were degraded much more rapidly than the OCFAs. The slow anaerobic degradation rate of OCFAs is likely linked to the initial concentration of the supplemented FAs, OCFA degraders abundance and the presence of propionic (C_3) acid as one of the intermediates. The results showed that anaerobic FAs oxidation by VFA-starved AD communities generates a variety of intermediate metabolites, which presumably affect the coexistence of active syntrophic bacteria and methanogenic archaea whose presence was detected by BONCAT.

4 Cell surface labelling and enrichment of active cells via biorthogonal non-canonical amino acid tagging and affinity-based cell separation

4.1 Introduction

The previous chapter described the degradation of medium-chain fatty acids to propionic and acetic acids, which can eventually be converted to methane. The observed catabolic profile might be mediated by a small number of active microorganisms that metabolise the compounds to support their survival and growth in AD systems. Unfortunately, marker gene profiling and metagenomic sequencing alone cannot provide information concerning which members of the community are active or the close associations that may form between syntrophic organisms in this anoxic habitat. Approaches that facilitate charting of process-targeted variation in microbial community activities are important for understanding how the microbiology of AD functions as a single biological machine.

The active microbes, as dynamic entities, react towards substrate availability or other changes in the environment by constantly adjusting their protein synthesis which will drive their cellular function. This postulation has been used by Dieterich et al. (2006) to introduce bio-orthogonal noncanonical amino acid tagging (BONCAT), a technique to label newly made proteins in system of interests via incorporation of noncanonical amino acid. These proteins can then be separated from their unlabelled counterparts via conjugation to an affinity tag and subsequently identified and quantified by liquid chromatography-tandem mass

spectrometry (LC-MS/MS). This technique has been applied successfully to study metabolically active microbes in various environmental samples (Hatzenpichler et al., 2014; Samo et al., 2014; Hatzenpichler et al., 2016; Leizeaga et al., 2017; Couradeau et al., 2019). This method has enabled the development of a new approach which is compatible with downstream genomic analyses to identify metabolically active microbes in AD via protein labelling. It enables cell-selective separation from the rest of the community via copper-catalysed click chemistry (Dieterich et al., 2007).

In this chapter, a recently developed protein labelling technique that could be adapted for detecting intact and active microbial cells in mixed pure cultures and AD samples is introduced. The enrichment of the intact labelled cells via affinity-based cell sorting (ABCS) can facilitate the identification of the active cells. It is applied first in a simpler system for downstream genomic analyses to demonstrate the specificity and sensitivity of BONCAT-ABCS in a mixed microbial system before being used to separate labelled cells from the complex mixture found in AD.

4.2 Experimental Design

The experimental approach (Figure 4.1) was designed to demonstrate cell surface labeling and enrichment of active *Escherichia coli* cells during mixed *E. coli* strains glucose-lactose diauxic growth. The *E. coli* glucose-lactose diauxic experiment of Mostovenko et. al (2011) was reproduced using *E. coli* MV1300 (MG1655 Δ *lacZYA*; kanamycin resistance) and *E. coli* MV1717 (MG1655 *lac+* and plasmid-encoded, chloramphenicol resistance), provided by Dr. Marjan van der Woude (HYMS, University of York). The growth conditions and the bioreactor set-up and operation can be seen in sections 2.1.1 and 2.2.1, respectively.
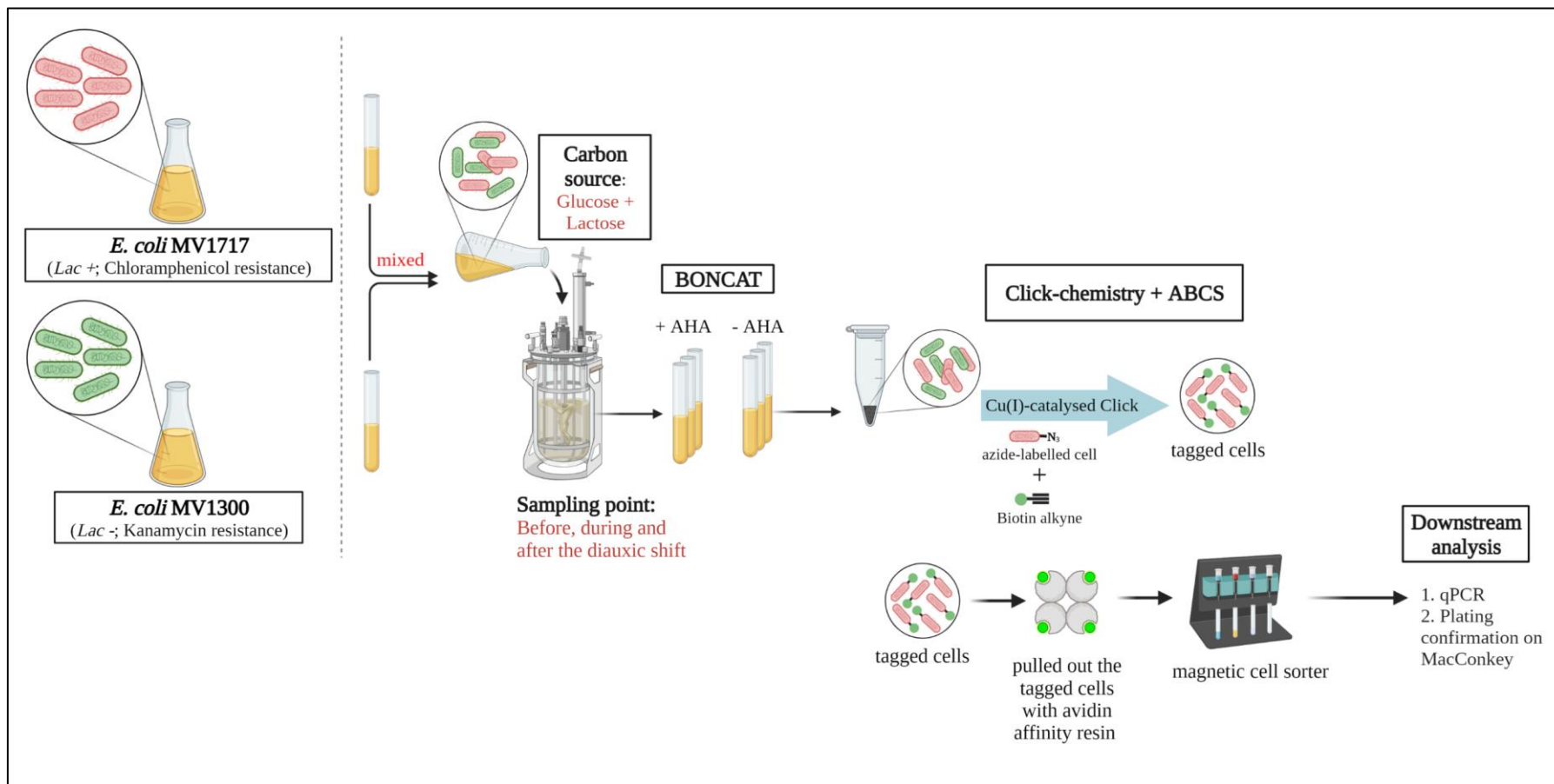
Figure 4.1 BONCAT-ABCS workflow for cell surface labelling and enrichment of active cells. Experimental design for cell surface labelling and enrichment of active *E. coli* cells during mixed *E. coli* strain glucose-lactose diauxic growth via BONCAT-ABCS. Created using Biorender.com.

2 mL samples were collected every 30 minutes before and after diauxie and every 10 minutes near and during the diauxic shift, as described in Mostovenko et al (2011), for monitoring the growth of the cells. *E. coli* cell growth was measured by assessing OD600 as described in section 2.3.5. The concentrations of glucose and lactose were assayed using enzymatic kits (MAK263, Sigma-Aldrich and K624, BioVision, respectively) as described in sections 2.3.3 – 2.3.4. Aliquots of cells were also cultured on MacConkey agar and incubated at 37 $^{\circ}$C overnight for differentiation and enumeration of lactose and non-lactose fermenting strains.

For BONCAT, 5 mL samples were transferred aseptically into a 15 mL Falcon tube. Samples were supplemented with L-azidohomoalanine (AHA; Click Chemistry Tools, Scottsdale, AZ, USA) in ultrapure water (0.2 µm filter sterilised) to 1 mM final concentration. After incubation (37 $^{\circ}$C, 15 min), samples were centrifuged (6,000 rpm, 15 min, 4 $^{\circ}$C; Eppendorf 5810R). Supernatants were discarded, pellets containing AHA-tagged cells were fixed with 3 mL of ethanol:PBS (1:1 v/v) and stored at -20 $^{\circ}$C before further processing. Click-labelling of chemically fixed microbial cells and BONCAT-labelled biomass recovery was essentially as described in sections 2.4.1 – 2.4.2. Genomic DNA was extracted from BONCAT-labelled biomass using a DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) as described in section 2.5.1.1. PCR and qPCR set-up were performed as described in section 2.5.2.

4.3 Results and Discussion

4.3.1 Mixed *E. coli* strains glucose-lactose diauxic growth

The glucose-lactose diauxie, a classic *E. coli* experiment, using mixed *E. coli* strains was recreated to show that BONCAT-ABCS works on a simple mixed

system. The strain MV1717 can grow on lactose (*lac+*), while MV1300 cannot utilise lactose (*lac-*) as it is missing the *lacY* gene that encodes lactose permease, a transporter that assists to pump lactose into the cells. This characteristic was confirmed by their growth on MacConkey agar (Figure 4.2A), where MV1717 (*lac+*) colonies grow pink and MV1300 (*lac-*) colonies grow colourless (white). Moreover, each strain contains chromosome coding an antibiotic resistance gene for chloramphenicol (MV1717) or kanamycin (MV1300) (Figure 4.2B-C), which were used as strain markers later on in qPCR analysis to determine the active strain in this study.
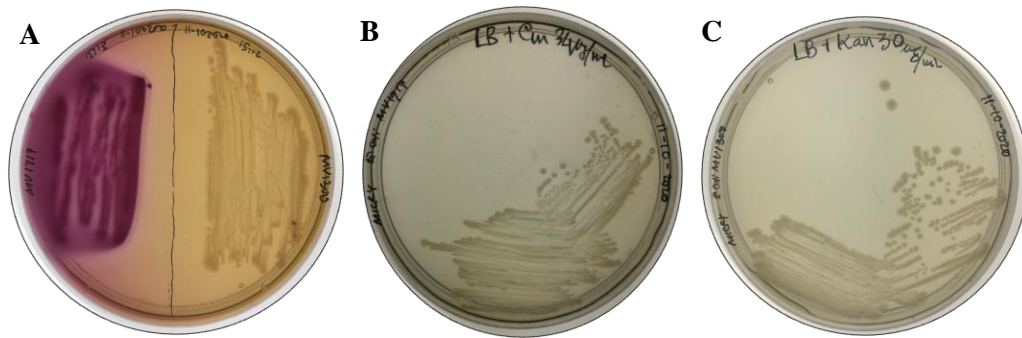
Figure 4.2 *E. coli* strains growth on a differential media and a rich media with antibiotic. A) *E. coli* strain MV1717 (left) and MV1300 (right) growth on MacConkey agar, B) strain MV1717 growth on LB agar with 34 µg/mL chloramphenicol and C) strain MV1300 growth on LB agar with 30 µg/mL kanamycin.

Both strains were grown separately for 18 h, until OD600 was ~1.4 (equivalent to ~1.1×10^9 cells/mL), and mixed in 1:1 ratio (v/v) prior to inoculation. Calculated OD600 of inoculum based on diluted samples was 6.82. The growth rate, lactose, and glucose concentrations in the experimental set-up allowed the accurate establishment of the onset of diauxic growth (Figure 4.3). Diauxie began when the cell suspension reached OD600 of ~ 0.5 or a density of approximately 4×10^8

cells/mL (Brown, 2020) and was indicated by a 20 – 30 minute plateau in the growth curve (Fig. 4.3). This is reproducible in each experiment (OD600 of 0.52, 0.55, 0.59) and Mostovenko et al. (2011). The onset of diauxie corresponds to the medium's glucose exhaustion. Lactose is depleted after ~250 minutes of diauxic shift and the growth reached stationary phase when OD600 ~2. OD600 0.5 – 0.6 was then used as a predictor during the experiment to optimise the sampling of the culture for active cell identification before, during and after the diauxic shift.
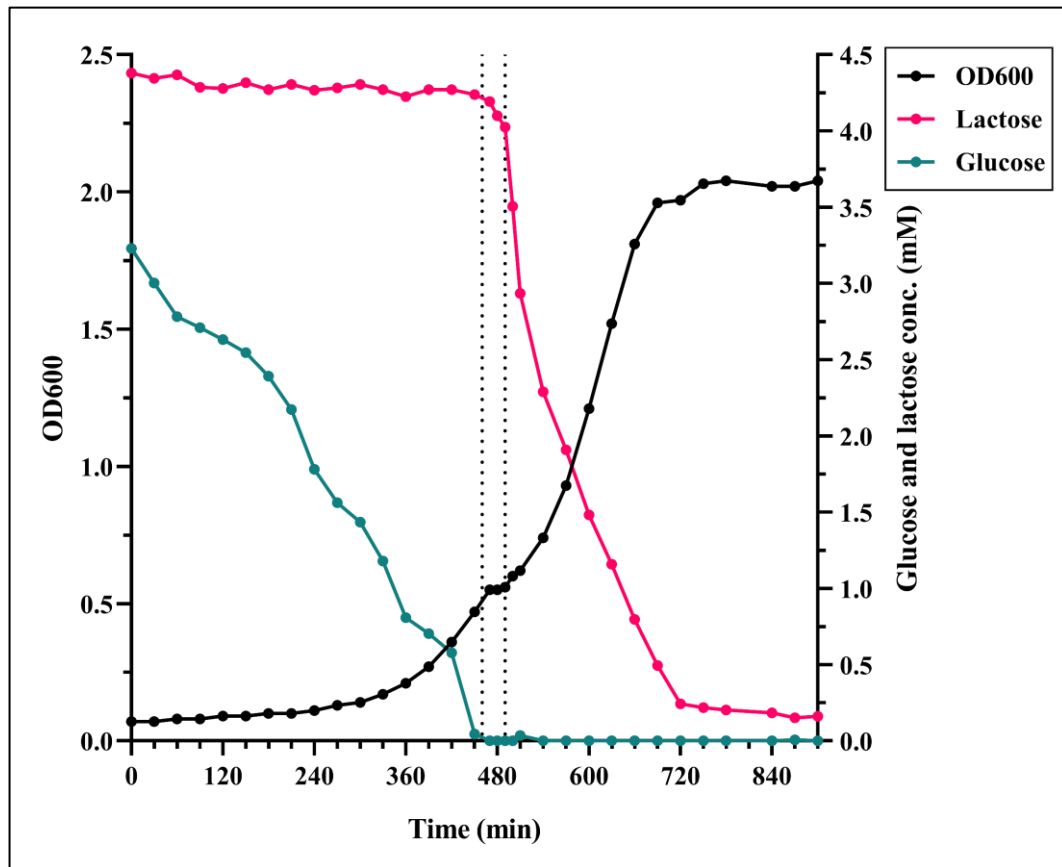
Figure 4.3 Mixed *E. coli* strains diauxic growth profile on glucose and lactose. Dashed lines indicated the onset of diauxic shift.

4.3.2 Cell surface labelling and enrichment of active *Escherichia coli* cells via BONCAT-ABCS

Based on the growth profile (Figure 4.3), aliquots in exponential growth before the diauxic shift (BD; 420 min), during the diauxic shift (DI; 470, DII; 480 and DII; 490 min) and exponential growth after the diauxic shift (AD; 600 min) were taken for active cells labelling via BONCAT. To use BONCAT for cell-selective analysis, the samples were supplemented with the presence or absence of 1 mM final concentration of the synthetic amino acid azidohomoalanine (AHA) for temporal selectivity of active cells protein labelling. This concentration was selected based on the premise that *E. coli* protein synthesis and degradation rates are not perturbed when supplemented with 1 mM AHA (Bagert et al., 2014; Hatzenpichler et al., 2014). A short (15 min) AHA incubation was performed, which is equivalent to 16 – 19 % of the generation time respective to a doubling time of 77 minutes in the glucose growth (between 360 and 450 minutes) and 94 minutes (570 - 660 min) in lactose growth. This incubation period was used to reduce excessive substitution of L-Methionine with AHA which can impede the cellular machinery efficiency (Hatzenpichler et al., 2014). AHA-labelled cells were then tagged with 3 µM biotin-alkyne via copper-catalysed click chemistry, as proven by Link and Tirrell (2003) to tag *E. coli* OmpC outer membrane protein, and concentrated on magnetic neutravidin beads as a means of affinity-based cell separation (ABCS). Neutravidin magnetic beads are microparticles (nominally 1 µm in diameter) with highly active neutravidin bound to the surface which can bind up to four biotinylated-alkyne groups with high affinity and selectivity. By estimating total surface area (3.14 μm^2) of the 1 µm bead, and dividing it by the top/bottom (0.785 μm^2) or lateral (3.14

μm^2) surface area of *E. coli*, it was found that the possible number of bound *E. coli* is 1 – 4 cells/bead. So, the amount of beads added 1:25 ratio of biomass for capturing the AHA-labelled cells is quite reasonable. The cells recovered were subjected to genomic DNA extraction for qPCR analysis to determine the active *E. coli* strain.

Prior to the qPCR assay, amplification specificity confirmation was performed and used to establish the amplification efficiencies of the primer sets used in this study (for the details of the primers see section 2.5.2.1). Three primer sets specific to the 30S ribosomal protein S17 gene (*rpsQ*), the chloramphenicol acetyltransferase gene (*catA1*; Pholwat et al., 2019) and the aminoglycoside 3'-phosphotransferase gene (*nptII*; Ullmann et al., 2019) were used to represent the housekeeping, chloramphenicol and kanamycin resistance genes, respectively. The amplification specificity was checked by both melting curve analysis and gel electrophoresis. All genes show a single melting peak and each PCR product also generated prominent bands with expected amplicon sizes (*rpsQ* 97bp, *nptII* 129 bp, and *catA1* 110 bp) in the gel electrophoresis analysis (Figure 4.4 – 4.5). These results confirmed that the primer sets used in this study did not generate non-specific PCR products during amplification. The primers were also used to generate qPCR standard curves (Figure 4.6), ranging from 0.09 – 6.02 ng/µL. All curves showed high linearity with correlation coefficient (R^2) > 0.999. Primer amplification efficiencies of 97.18 %, 86.65 % and 93.72 % were obtained from the slopes of their corresponding standard curves (-3.3913, -3.6897 and -3.4822) for *rpsQ*, *nptII* and *catA1*, respectively. The standard curves constructed (Fig. 4.6) were used to quantify the copy number of chloramphenicol and kanamycin resistance

gene that has been normalised to the housekeeping gene as a proxy of strain MV1717 and MV1300 abundance in each sample.
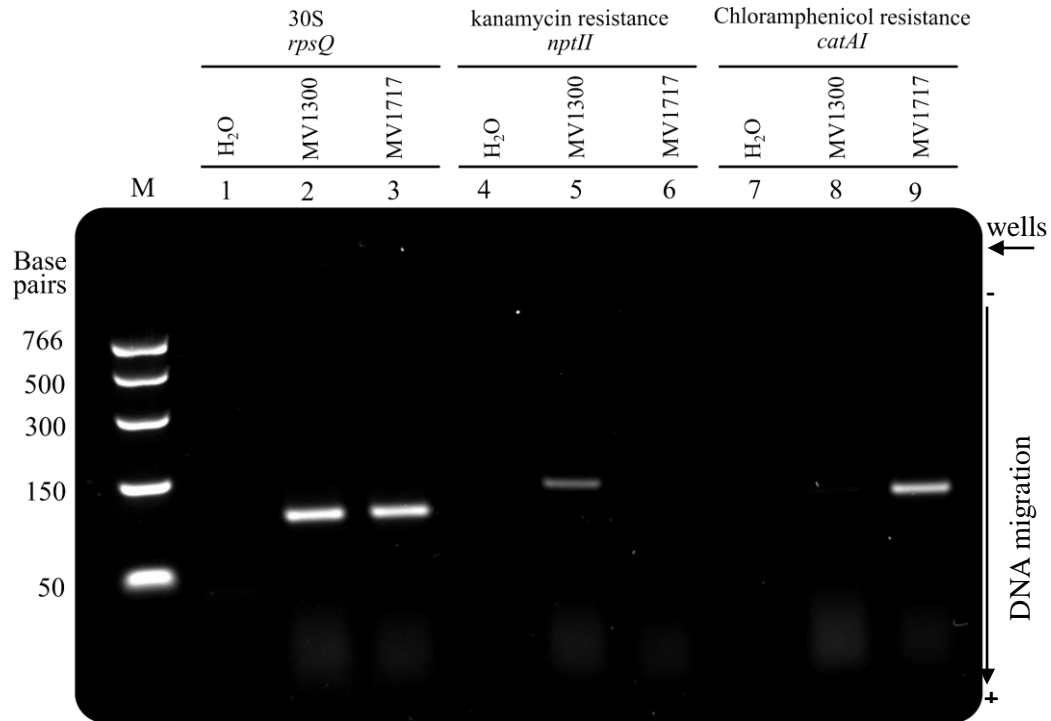
Figure 4.4 The primer amplification specificity confirmation. Amplification specificity confirmation of each primer PCR product run on a 3% agarose gel in 1x TBE, 5 volt/cm, stained with SYBRSafe DNA gel stain. M = low molecular weight DNA ladder (NEB); 1 = negative control *rpsQ* (water); 2 = *rpsQ* – MV1300; 3 = *rpsQ* – MV1717; 4 = negative control *nptII* (water); 5 = *nptII* – MV1300; 6 = *nptII* – MV1717; 7 = negative control *catA1* (water); 8 = *catA1* – MV1300; 9 = *catA1* – MV1717. The position of the wells and direction of DNA migration is noted.
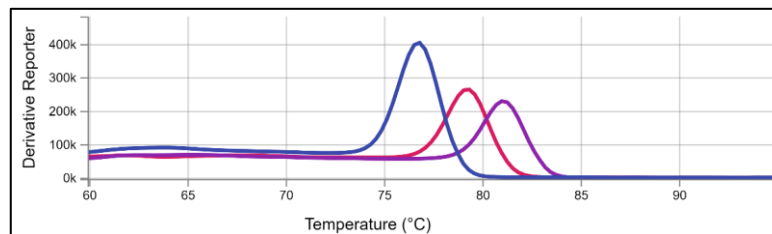
Figure 4.5 Derivative melting curve analysis of each primer PCR product shows primer specificity. A single peak indicates a single PCR-product. Colours denote *rpsQ* (red), *nptII* (purple), and *catA1* (blue). The melting curve analysis with a temperature gradient of 0.1 °C/s from 60 to 95 °C was performed using the Applied Biosystems® QuantStudio® 3 Real-Time PCR System.
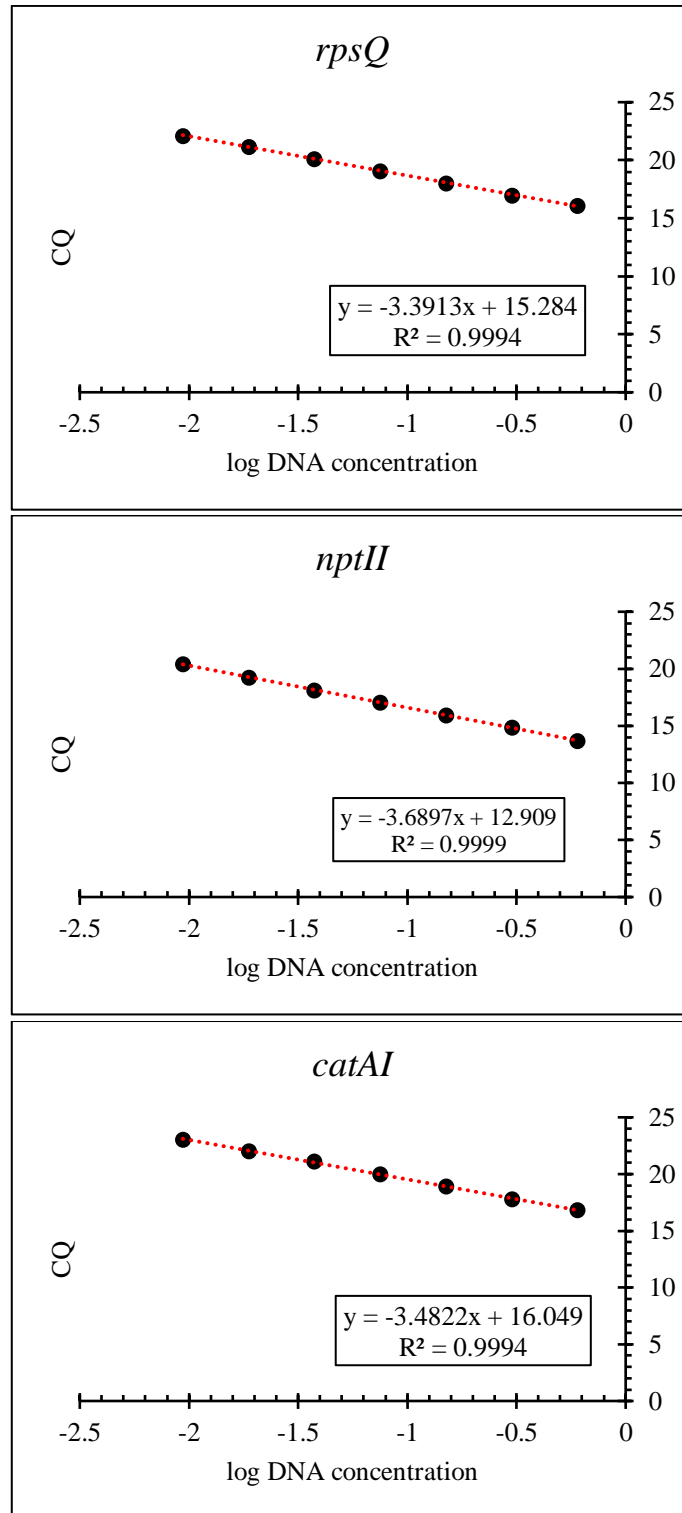
Figure 4.6 Evaluation of qPCR primer efficiencies. The amplification efficiency for each primer quantitation cycle (CQ) and the logarithm of the initial DNA concentrations were plotted to calculate the slope of each primer pair. Standard curves were generated from at least seven dilution points for each primer pair. qPCR reactions for each sample were run in triplicate. Amplification efficiencies were calculated according to the equation $E = 10^{(-1/slope)}$ (Rasmussen, 2001).

126

During exponential growth before the diauxic shift, while using glucose, both MV1300 and MV1717 strains grew and increased in cell density in approximately equal numbers, as shown by kanamycin and chloramphenicol gene abundances for both BONCAT and control samples. Strain MV1717 was slightly more abundant than MV1300 as demonstrated by BONCAT-ABCS, although not significantly different (Fig. 4.7), confirmed by the number of pink (*lac+*) and white (*lac-*) colonies grown on MacConkey agar (Figure 4.8).

In the course of the diauxic shift, MV1717 starts to express the genes required for lactose metabolism while MV1300 may express genes to support stationary phase and both strains stop increasing in cell density (Figure 4.3). Chloramphenicol gene abundance was higher than the abundance of the kanamycin gene in BONCAT samples, compared to control samples where both antibiotic genes were detected in apparently equal abundance at the beginning of diauxic shift. BONCAT-ABCS revealed that the MV1717 strain was becoming more active in the course of diauxic shift, while the MV1300 strain remained persitent during the diauxic shift, which was confirmed by the enumeration of lactose and non-lactose fermenting strains on MacConkey agar (Figure 4.8A and C-E). The results of three consecutive samples (DI, DII, DIII) analysed via BONCAT-ABCS demonstrate the specificity and the robustness of this method.
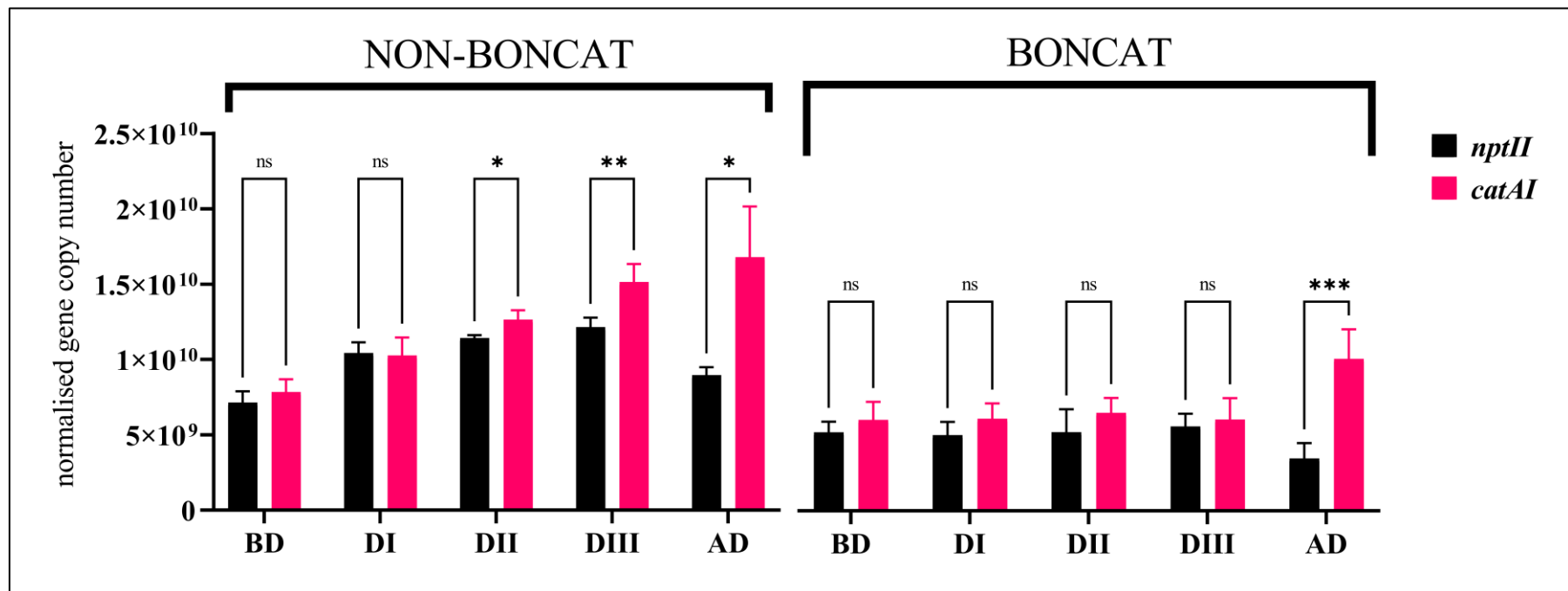
Figure 4.7 Quantification of MV1717 and MV1300 abundance between non-AHA-labelled (NON-BONCAT) versus AHA-labelled (BONCAT) samples during the diauxic growth. The *nptII* and *catA1* gene were used as a marker of each strain normalised to *rpsQ* (the housekeeping gene). BD = samples in exponential growth before the diauxic shift (420 min); DI - DIII = samples during the diauxic shift (470, 480 and 490 min, respectively); AD = samples during exponential growth after the diauxic shift (600 min). Bars represent the standard deviation of six replicates (n=6). The significance of differences was analysed by two-way ANOVA test (**** $P < 0.0001$; ns, not significant) and performed using GraphPad Prism software version 9.0.2.

In exponential growth after diauxic shift where lactose is the sole carbon source, only MV1717 can grow and the cell density increases, yet the MV1300 strain is still present (Figure 4.7 and 4.8A and F). Chloramphenicol gene abundance was significantly high for BONCAT samples with low kanamycin gene abundance observed. This is supported by notably more pink (*lac+*) than white (*lac-*) colonies grown on MacConkey agar (Figure 4.8). A substantial amount of active MV1300 strain is still present in the sample because these survivors may use the nutrients released by the dead cells and brings the strain to a state of long-term stationary phase (Pletnev et al., 2015).
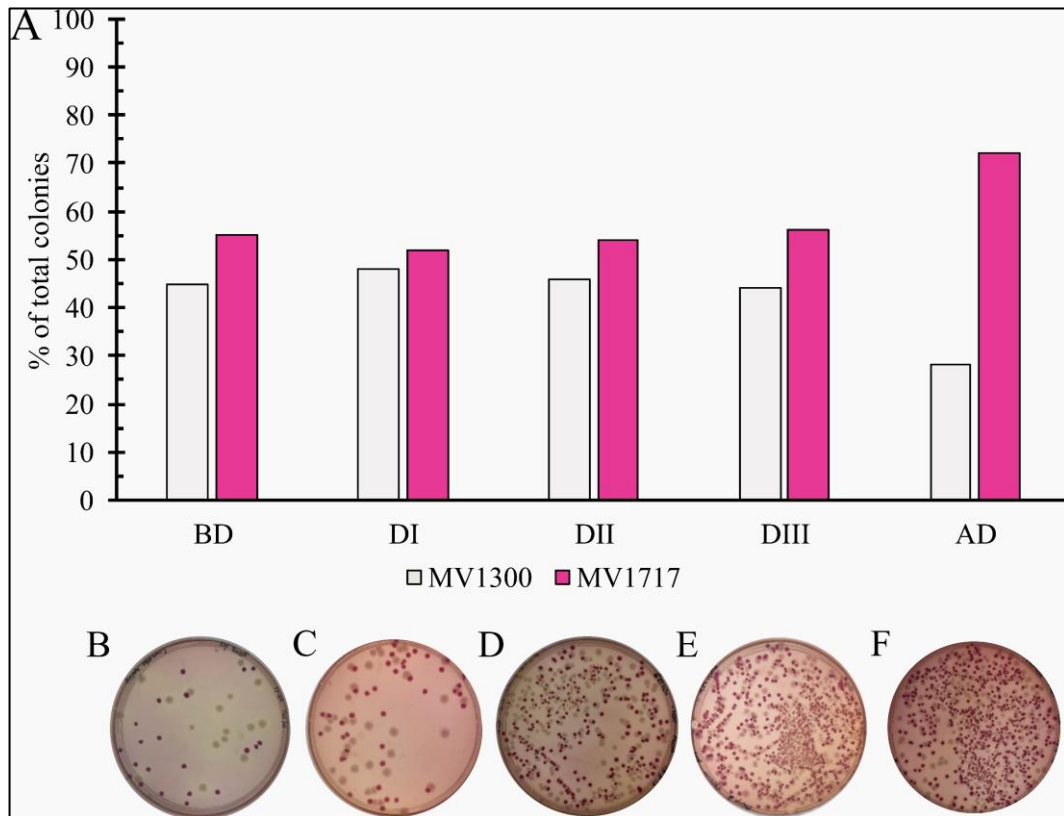
Figure 4.8 Colony growth (A) and counts of lactose and non-lactose fermenting strain on MacConkey agar (B – F). BD = samples in exponential growth before the diauxic shift (420 min); DI - DIII = samples during the diauxic shift (470, 480 and 490 min, respectively); AD = samples during exponential growth after the diauxic shift (600 min).

4.3.3 Affinity-based cell separation enriches active cell populations in AD sample

To establish whether affinity tags could be used to separate labelled cells from the complex mixture found in AD, a biomass sample from 12 h post octanoic acid addition that had been incubated in the presence of 1 mM AHA was used and supplied the click cocktail with a 50:50 mix of biotin and Alexa-488 alkynes (Appendix A). The biotin-tagged cell population was isolated from the community using 10 µL (~ 100 µg) Neutravidin magnetic beads. 20% glycerol was used to increase sample density to improve magnetic separation as it is cheap and biologically non-invasive (Appendix B; Volk and Kähler, 2018). Fig. 4.9 shows that the BONCAT-labelled cells were enriched on the magnetic beads as a means of affinity-based cell separation (ABCS). The result suggests that both biotin and Alexa-488 moieties were incorporated into some cells and that these were selectively enriched onto Neutravidin beads, consistent with the labelling reaction resulting in a fraction of nascent proteins locating to the surface of metabolically active cells. The successful enrichment of AHA-based BONCAT followed by ABCS indicates that this is a viable approach for isolating populations of microbes actively responding to substrate availability in AD systems for further analysis.

DAPI
(DNA)

BONCAT
(Active cells)
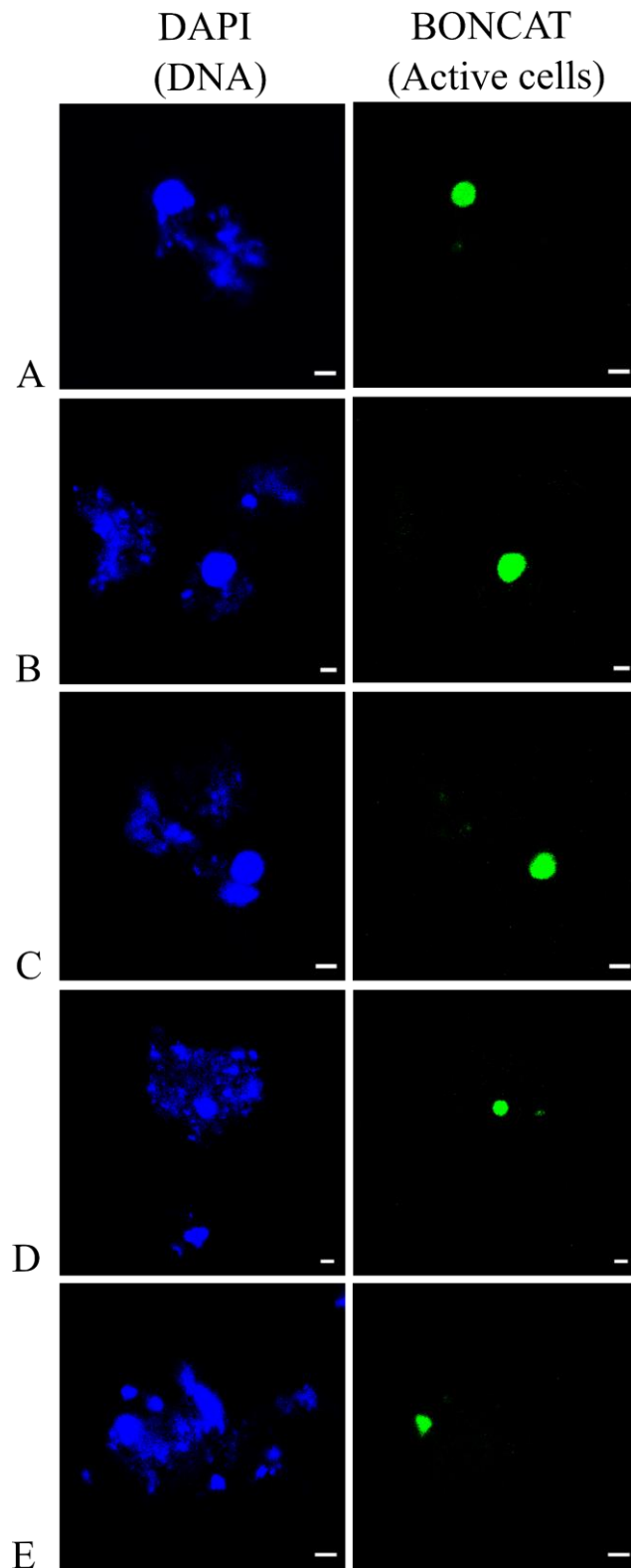
Figure 4.9 Visualisation of translationally active cells immobilised on Neutravidin beads. The visualisation was done by combining BONCAT with affinity-based cell sorting (ABCS). Each row (A-E) shows a separate field of view. Labelled cells are shown in green and DAPI staining of DNA in blue. Scale bars indicate 2 µm in all images.

4.4 Conclusion

These results suggest that biotin-alkyne moiety was incorporated into some cells and that these were selectively enriched onto neutravidin beads, consistent with the labelling reaction resulting in a fraction of nascent proteins locating to the surface of metabolically active *E. coli* cells. The BONCAT-ABCS technique was applied successfully and displayed the specificity and sensitivity to enrich the active cells during the diauxic growth of mixed *E. coli* strains used in this study. The successful enrichment of AHA-based BONCAT followed by ABCS indicates that this is a viable approach for isolating populations of microbes actively responding to substrate availability in mixed systems.

5 Bioorthogonal non-canonical amino acid tagging and affinity-based cell separation provide insight into octanoic acid degrading microbial communities

5.1 Introduction

Our understanding of AD microbial communities relies strongly on 16S rRNA gene profiling and metagenomic sequencing which can reveal phylogenetic diversity (Wirth et al., 2012; Kirkeegard et al., 2017; Hassa et al., 2018) but does not provide information concerning microbial activity or the state of cells in the system. Metagenomic data can provide information on organisms that are present, but not necessarily active, within AD systems. There are a range of approaches for targeting active microbes in AD samples including micro-autoradiography (MAR; Lee et al., 1999), secondary ion mass spectroscopy (SIMS and nano-SIMS; Musat et al., 2008; Orphan et al., 2009; Eichorst et al., 2015; Berry et al., 2015), and Raman micro-spectroscopy (Raman; Huang et al., 2007; Eichorst et al., 2015; Berry et al., 2015). While these methods are effective at identifying active microbes, they are often destructive and cannot be easily combined with downstream genomic analyses (Hatzenpichler et al., 2020).

In the past few years a number of groups have used bioorthogonal non-canonical amino acid tagging (BONCAT) to identify metabolically active microbes in pure culture and environmental samples, such as marine sediments, pond sediments, and soil (Hatzenpichler et al., 2016; Leizeaga et al., 2017; Hatzenpichler et al., 2014; Samo et al., 2014; Couradeau et al., 2019). This technique has not been used to track the activity or monitor microbial community changes in anaerobic digestion.

BONCAT makes use of the synthetic amino acid analogues L-azidohomoalanine (AHA) or L-homoproparglycine (HPG) as surrogates of L-methionine (Kiick et al., 2002; Hatzenpichler et al., 2014; Couradeau et al., 2019) to label L-methionine-containing proteins in any organism where protein synthesis is occurring. BONCAT has been shown to be effective in a broad range of microorganisms and is seemingly independent of metabolism (Hatzenpichler et al., 2016; Samo et al., 2014). Labelled proteins from active cells can be visualized using fluorescence microscopy after attaching fluorescent dyes via Cu(I)-catalysed click chemistry (Hatzenpichler et al., 2016; Dieterich et al., 2007) to the incorporated bioorthogonal amino acids. Click chemistry can be used to attach alternative reagents such as biotin alkyne (Dieterich et al., 2007) allowing tagged proteins to be isolated via avidin resins (Dieterich et al., 2007; Dieterich et al., 2006; Pezzi et al., 2018). Compared to the previously mentioned methods BONCAT is fast (Couradeau et al., 2019) and uses small amounts of material (Hatzenpichler et al., 2016; Couradeau et al., 2019), facilitating the collection of multiple samples over time and making it an appealing option for tracking microbial community changes or activity in anaerobic digestion.

The previous chapter describes the application of temporally-selective cell labelling and enrichment using BONCAT and affinity-based cell separation (ABCS) in a simple system, using mixed pure cultures, for downstream genomic analyses. Here, the successful application of BONCAT combined with affinity-based cell separation to facilitate the specific enrichment of metabolically active microbial populations in AD-derived sludge during octanoic acid degradation is reported.

5.2 Experimental Design

The experimental approach (Figure 5.1) was designed to demonstrate the applicability of BONCAT-ABCS as a tool for identifying active microbes in complex AD microbial communities. It was used to investigate the activities of VFA-starved microbial communities in the presence of octanoic acid compared to nutrient rich synthetic feed (positive control) and water (negative control). The composition of the synthetic feed is described in Table 5.1.

Table 5.1 Composition of the synthetic feed (Tao et al., 2019)

Constituents	Conc. (g/L)	Trace Elements	Conc. (mg/L)
KH_2PO_4	0.31	Fe ($FeCl_2 \cdot 4H_2O$)	10.0
Na_2HPO_4	2.47	Ni ($NiCl_2 \cdot 6H_2O$)	1.0
$MgCl_2 \cdot 6H_2O$	0.15	Se (Na_2SeO_3)	0.2
$CaCl_2 \cdot 2H_2O$	0.11	W ($Na_2WO_4 \cdot H_2O$)	0.2
Urea	1.20	Mo (($NH_4)_6Mo_7O_{24} \cdot 4H_2O$)	0.2
Yeast extract	3.90	Co ($CoCl_2 \cdot 6H_2O$)	1.0
Sucrose	23.30	Al ($AlCl_3 \cdot 6H_2O$)	0.1
		Zn ($ZnCl_2$)	0.1
		Mn ($MnCl_2 \cdot 4H_2O$)	0.1
		B (H_3BO_3)	0.1
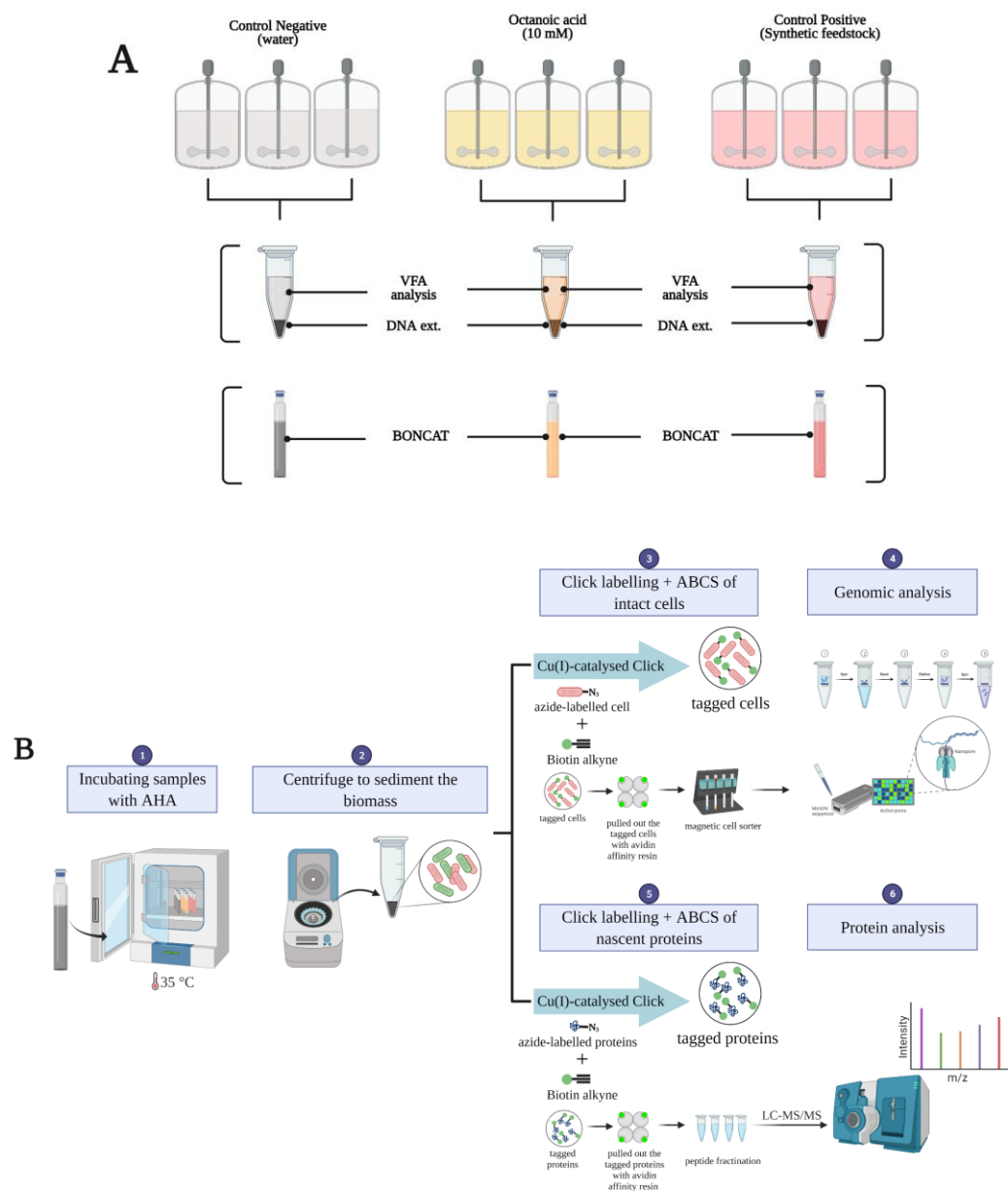
Figure 5.1 Experimental design for active cell labelling and enrichment of active AD microbial communities via BONCAT-ABCS. (A) Incubations were done in triplicate and samples were subjected to VFA measurement and (B) BONCAT-ABCS of intact cells and proteins for each treatment. Created using Biorender.com.

'Starved' inocula were generated using material collected from a process-scale (1,858 m^3) AD system at Yorkshire Water's Naburn site, York, United Kingdom (53°54'50.5"N 1°05'04.6"W) as described in section 2.1.2. Samples were incubated

in stirred 5 L reactors (section 2.2.2; Figure 2.1) at 35 °C until VFA content was undetectable by GC-FID (HP 5890 series II) measurements (< 0.05 mM; see section 2.3.1). VFA-starved communities were then used as the starting point for this experiment. Communities were fed with a single addition of 8 mL of 6.310 M octanoic acid (~10 mM final concentration), synthetic feed or water. Incubations were carried out in triplicate (Fig. 5.1A) and six time-points were subjected to VFA measurement (section 2.3.1) and BONCAT-ABCS (section 2.4.1 and 2.4.3) for each treatment. Genomic DNA was extracted from recovered biomass using a Qiagen PowerSoil DNA Extraction Kit (section 2.5.1.2) and subjected to metagenomic sequencing (section 2.5.3). To link the observed metabolic activity to specific microbial taxa, labelled proteins extracted from BONCAT-ABCS samples (section 2.4.4) were subjected to protein identification (section 2.5.4). Metagenomic and protein analysis were done as described in section 2.6 and 2.7, respectively. pH was measured using pH indicator test paper (Whatman 2600103A pH 6.4 – 8.0 range, GE Helathcare).

5.3 Results and Discussion

BONCAT-ABCS based protocols have been successfully employed for the study of enriched nascent proteins in mammalian and *Leishmania* cells (Dieterich et al., 2006; Bagert et al., 2014; Kalesh and Denny, 2019). However, it has not yet been tested for enriching whole cells within complex microbial communities, due to the challenges presented, such as AHA uptake and variation in growth stages of microbes in the community (Valentini et al., 2020). The activity-based labelling method developed in this study overcomes these challenges and allows enrichment of active microbial aggregates in situ that express the probe-containing proteins on

their cell membrane, as demonstrated in Chapter 4. For this activity-based labelling method, VFA-starved AD communities were exposed to a specific substrate to stimulate expression of AHA-contained proteins by metabolically active microorganisms. The BONCAT-ABCS was combined with subsequent metagenomic and proteomic analyses to directly link the function of a population to its identity *in situ*, which helped to identify microorganisms capable of degrading octanoic acid.

VFA-starved anaerobic sludge was subjected to different substrate amendment with the initial pH of 7.1 ± 0.1 (Fig.5.2). The pH during AD of nutrient rich synthetic feed or water were consistent (pH 7.1 ± 0.1), while octanoic acid addition slightly lowered the pH to 6.9 ± 0.1. The pH observed in this study was still in the optimum pH range in AD, 6.8 – 7.2 (Cioabla et al., 2012), and has no negative effect on AHA stability (azide reduction to amine if pH > 7.5; Hatzenpichler and Orphan, 2015).
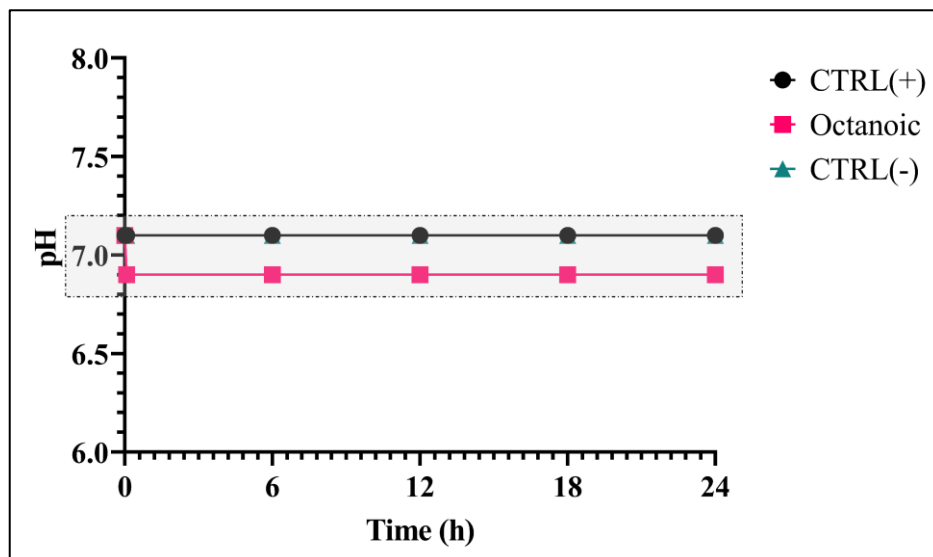
Figure 5.2 pH profile during AD of octanoic, nutrient rich synthetic feed (positive control) and water (negative control) by VFA-starved AD communities. Dashed grey box indicated the optimal pH range in AD (Cioabla et al., 2012).

5.3.1 Metagenomic data from BONCAT-ABCS reveal the extensive heterogeneity of translational activity in an AD microbial community

The VFA-starved microbial community metabolic activity was extrapolated from VFA profiles, which allowed us to observe the degradation of octanoic acid compared to the control treatments (Fig. 5.3). Octanoic acid was cleaved into shorter chain (hexanoic (C_6)) fatty acids (Fig. 5.3A-F). These shortened fatty acids then likely enter the β-oxidation pathway to produce acetic (C_2) acids which can ultimately be converted to methane. In contrast, nutrient rich synthetic feed (positive control) was metabolised to an array of VFAs. In this case, the major intermediate metabolised detected by the VFA-targeted GC method were acetic (C_2), propionic (C_3), *iso*-butyric (iC_4), butyric (C_4) and *iso*-pentanoic (iC_5) acids (Fig. 5.3G-L). The water fed VFA-starved community (negative control) showed no difference in VFA profile before or after addition. These results suggest that the type of substrate added affects the metabolic profile observed, which might indicate the activity of distinct active AD microbial taxa.
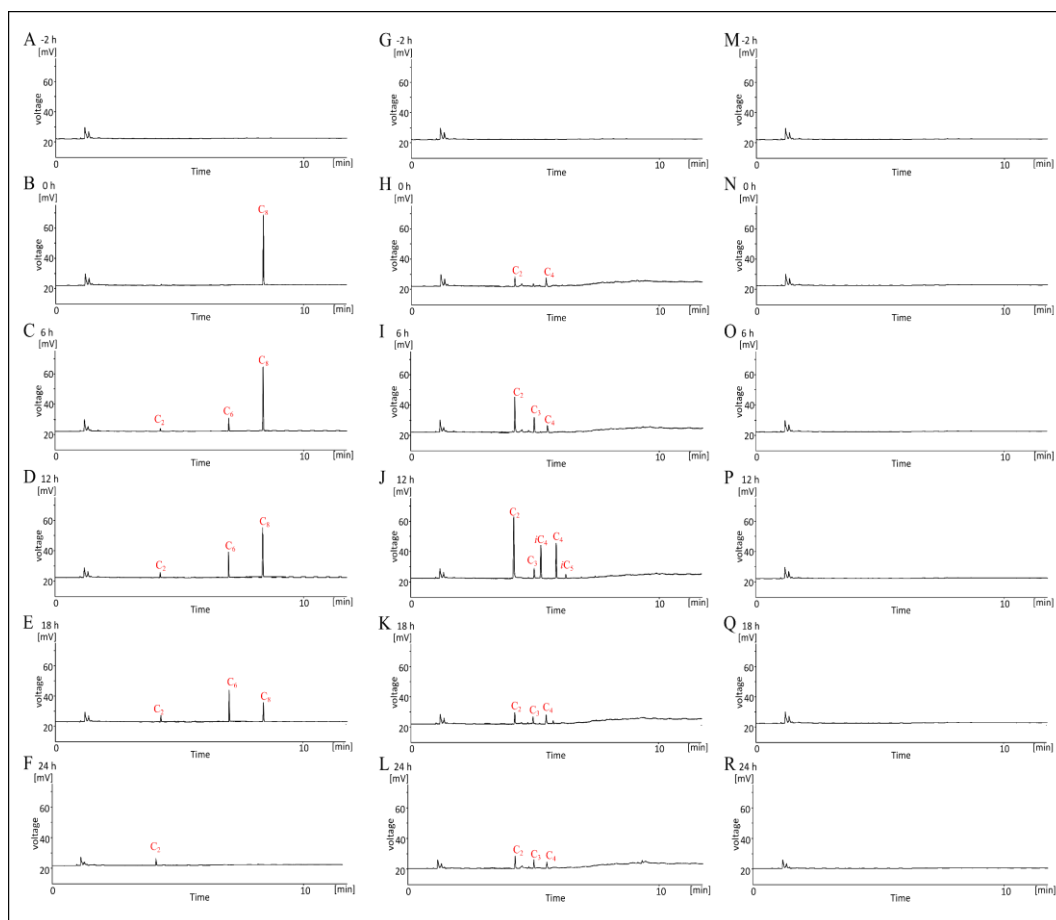
Figure 5.3 Comparison of volatile fatty acid (VFA) profiles during 24 hours of octanoic (C_8) acid (A-F), synthetic feed (G-L) and water (M-R) catabolism by AD-derived sludge. (A,G,M) Starved sludge before spike; (B) after C_8 spike; (C-F) C_8 after 6 – 24 h; (H) after synthetic feed spike; (I-L) synthetic feed after 6 – 24 h; (N) after water spike; (O-R) water after 6 – 24 h at 35 °C. C_2-acetic; C_3-propionic; iC_4-*iso*butyric; C_4-butyric; iC_5-*iso*pentanoic; C_6-hexanoic; C_8-octanoic acid.

Metagenomic analysis was applied to sorted and unsorted samples from each treatment to determine the composition of the microbial communities captured through BONCAT-ABCS. A total of 54 DNA samples of BONCAT-ABCS and non-BONCAT samples were extracted and sequenced (Appendix C), from three treatments. BONCAT-ABCS samples yielded significantly less DNA compared to the non-BONCAT samples, reflecting labelling of only a subpopulation of cells and indicating that specific substrate addition might only support a small population of

cells after starvation. DNA samples retrieved from BONCAT-ABCS samples were sequenced using Illumina short-reads which allows for more in-depth sequence coverage to be obtained.

Sequence data were analysed using the CLUSTard pipeline (https://github.com/ac1513/CLUSTard) to visualise abundance changes between samples and enable observation of fine-scale variation in bacterial populations (including genus-level resolution) based on GTDB taxonomy annotation (Parks et al., 2021). To characterise translationally/metabolically active subpopulations, sequence data from AHA-labelled samples were compared to their associated non-labelled/unsorted sample. The relative abundance proportions of the top-30 ranking taxa were used to compare sample replicates for each treatment (Appendix E – G). Replicates showed reproducible taxonomic profiles, so relative abundances were averaged for further analysis (Fig. 5.4 – 5.6).

In terms of genus-level resolution, each treatment had a high complexity AD microbiome (Fig. 5.4 – 5.6), which included syntrophic bacteria and methanogens (*e.g., Smithella, Syntrophomonas*, *Syntrophorhabdus, Methanothrix* and *Methanolinea*). In general, the microbial community structure retrieved from the BONCAT-ABCS in octanoic acid, synthetic feed, and water fed digesters displayed similarities among the most abundant community members relative to its original (unsorted) sample. The data suggest that a subset of the majority of genera identified by conventional metagenomic sequencing are translationally/metabolically active.
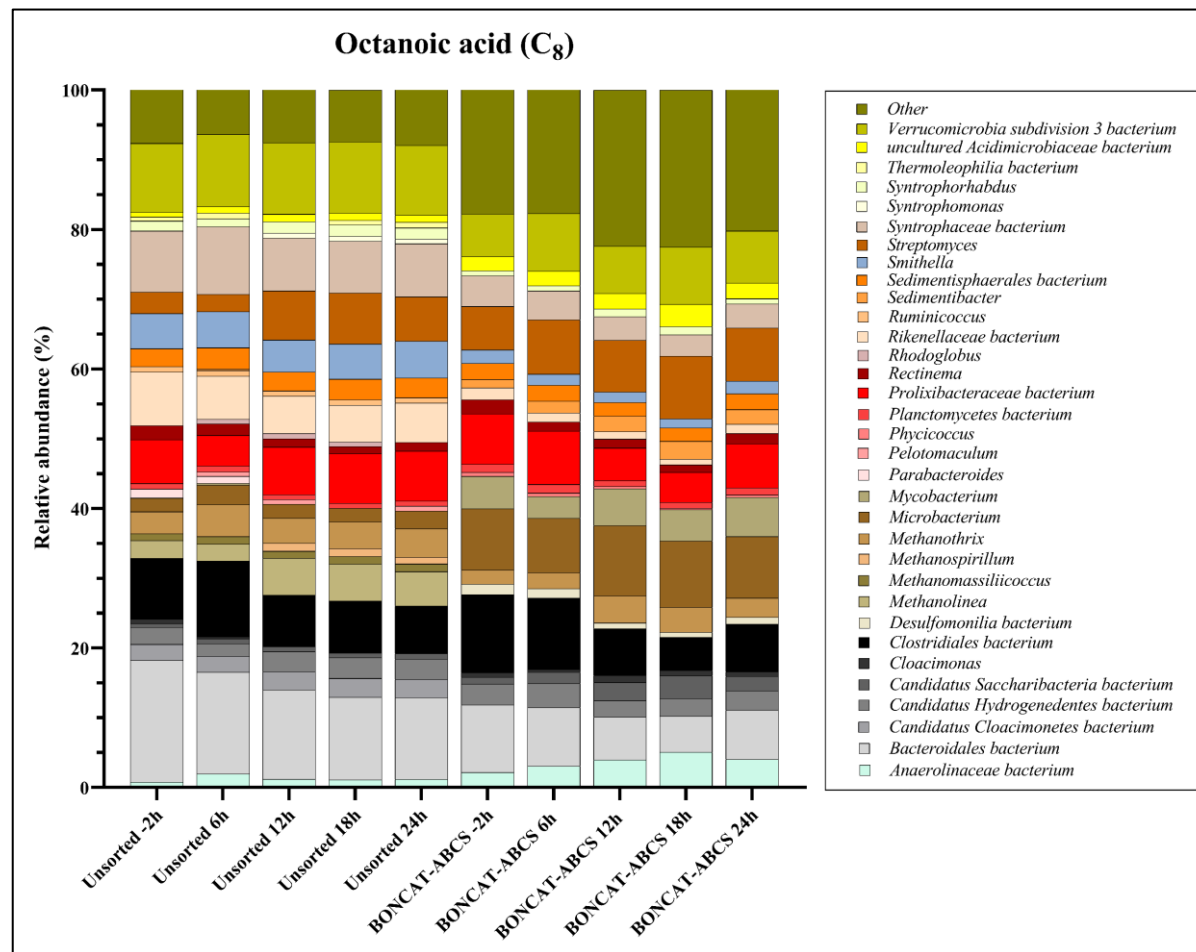
Figure 5.4 Relative abundance of top-30 most abundant genera based on metagenomic sequencing of sorted and unsorted samples retrieved from octanoic acid fed digesters.
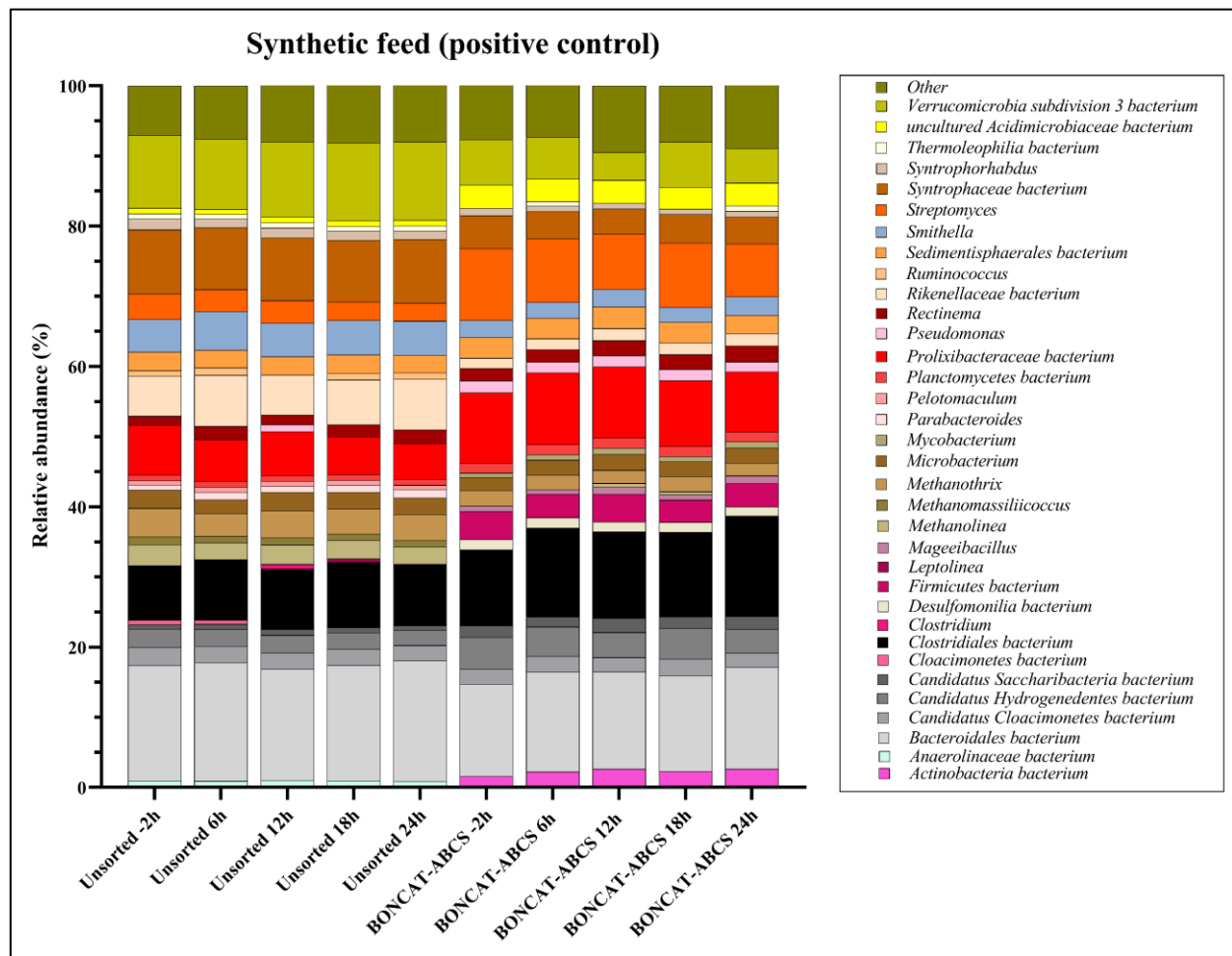
Figure 5.5 Relative abundance of top-30 most abundant genera based on metagenomic sequencing of sorted and unsorted samples retrieved from nutrient rich synthetic feed digesters.
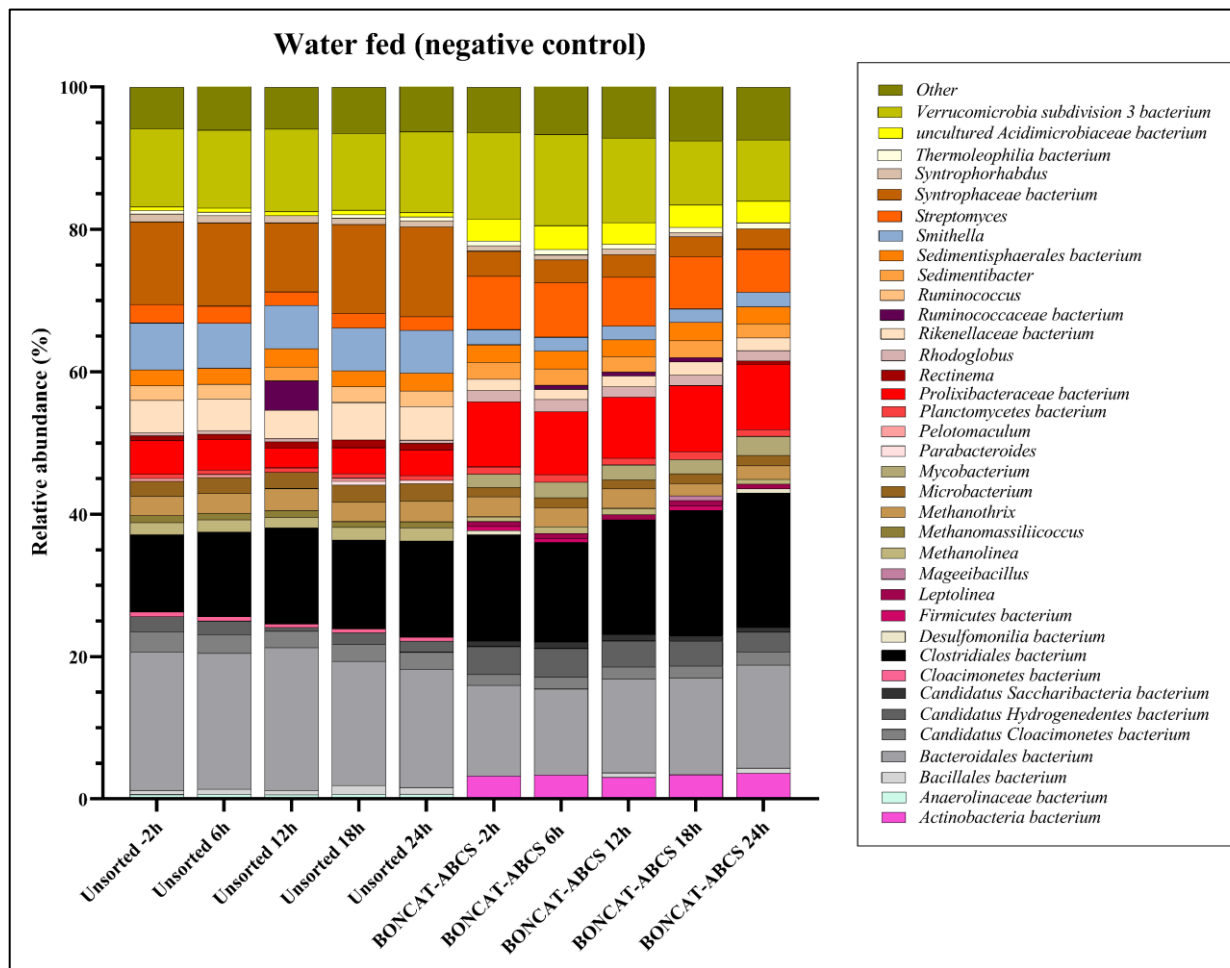
Figure 5.6 Relative abundance of top-30 most abundant genera based on metagenomic sequencing of sorted and unsorted samples retrieved from water fed digesters.

It is plausible that some of the active AD microbes in the presence of VFAs were also active in the absence of VFAs. To better observe causative relationship between changes in the relative abundances of translationally/metabolically active microbial taxa and the VFA degradation (Fig. 5.3), fold-change difference of enriched subpopulation in octanoic and synthetic feed (positive control) treatments compared to the water fed treatment (negative control) for each time-point were calculated (Fig. 5.7 and 5.8). Some genera present in taxa plots do not appear in fold-change plots (Fig. 5.4) because they only appeared in negative control samples and/or their relative abundance was less than 1%, although it should be highlighted that activity in these less abundant populations may also be determinants of octanoic and synthetic feed anaerobic degradation. For example, *Syntrophomonas* had an average relative abundance of ~ 0.1% throughout the course of octanoic acid degradation but showed high translational activity with 2 – 4 fold more abundant in octanoic samples compared to the negative control.

Generally, the relative abundance ranks of octanoic or synthetic feed compared to water fed fractions did not appreciably different (denoted by heatmaps in Fig. 5.7 and 5.8, respectively). One of the most abundant genera in all octanoic fed samples, *Microbacterium*, was consistently in larger relative abundance in the community structure retrieved from the BONCAT-ABCS samples, reflecting its active growth and highlighting its important role in octanoic acid degradation. However, this trend in agreement between relative abundance and fold-change did not always hold. For example, unclassified member of *Clostridiales* (Firm-04) was highly abundant in all octanoic timepoints (see heatmaps), but its fold difference was always larger in negative control (water fed) samples, indicating lower relative translational activity

than its co-colonising AD microbial community. This was also observed for unclassified members of *Bacteroidales* and *Verrucomicrobia subdivision 3* in octanoic and synthetic feed fractions, which were among the most abundant genera yet showed low translational activity.

In the octanoic acid fed system, the number of translationally active genera increased over time, reflecting the metabolic activity of these genera in the presence of octanoic acid and the resulting VFA intermediates during octanoic degradation (Fig 5.7). Before octanoic acid addition, A total of 23 differentially abundant genera were identified, 10 of which were showed high translational activity in the VFA-starved community. *Microbacterium*, *Rectinema* and an unclassified member of *Anaerolinaceae* (T78/UBA6107) showed the highest relative translational activity. Syntrophic bacteria, such as *Syntrophomonas* and unclassified members of *Syntrophaceae* (UBA2192/UBA8904) and *Desulfomonilia* (UBA1062), were also found to have high activity. Methanogens showed low translational activity as *Methanolinea* (hydrogenotrophic methanogen) and *Methanothrix* (acetoclastic methanogen) had 0.46% and 1.98% relative abundance at -2 h, respectively, but were 1.4-fold more prominent in the negative control. Six hours after octanoic acid addition, hexanoic and acetic acids were observed with a concomitant reduction in octanoic acid, more bacterial genera become translationally active, such as *Mageeibacillus*, *Streptomyces*, *Syntrophorhabdus*, and an unclassified member of *Firmicutes* (UBA3907/UBA5389). By 12 hours, methanogens (e.g., *Methanolinea* and *Methanothrix*) started to show high translational activity and kept increasing until 18 hours after addition of C_8. By 24 hours, the relative abundance fold-change of the translationally active bacteria and methanogens were decreased slightly as

neither octanoic nor hexanoic acids were detectable, although acetic acid was still present.

As a comparison, the enriched population retrieved from the synthetic feed digesters showed that the rich nutrient substrate induces more translationally active genera. 70% of identified highly abundant genera were translationally active which likely reflects the variability of metabolic activity driven by rich nutrient availability (Fig. 5.3). Bacteria belonging to *Rectinema* (UBA8932) and unclassified members of the *Firmicutes* (UBA3907/UBA4882/UBA5389) were highly active among the translationally active bacteria during 24 hours of synthetic feed degradation. *Methanothrix* and *Methanolinea* became highly translationally active in the 18 h and 24 h post-addition sample (Fig.5.8).

A. -2 h (before octanoic acid addition)

B. 6 h after octanoic acid addition

Figure 5.7 Fold-changes of active taxa in the octanoic acid compared to the water fed (negative control) fraction (A – E). Point colour indicates taxa that were increased (green) and decreased (pink) in relative abundance in the octanoic fraction, representing translationally/metabolically active genus. Heatmap sidebars represent square root transformed relative abundances.

Figure 5.7 Fold-changes of active taxa in the octanoic acid compared to the water fed (negative control) fraction (A – E). Point colour indicates taxa that were increased (green) and decreased (pink) in relative abundance in the octanoic fraction, representing translationally/metabolically active genus. Heatmap sidebars represent square root transformed relative abundances.

Figure 5.7 Fold-changes of active taxa in the octanoic acid compared to the water fed (negative control) fraction (A – E). Point colour indicates taxa that were increased (green) and decreased (pink) in relative abundance in the octanoic fraction, representing translationally/metabolically active genus. Heatmap sidebars represent square root transformed relative abundances.

Figure 5.8 Fold-changes of active taxa in the synthetic feed compared to the water fed (negative control) fraction (A – E). Point colour indicates taxa that were increased (green) and decreased (pink) in relative abundance in the synthetic feed fraction, representing translationally/metabolically active genus. Heatmap sidebars represent square root transformed relative abundances.

Figure 5.8 Fold-changes of active taxa in the synthetic feed compared to the water fed (negative control) fraction (A – E). Point colour indicates taxa that were increased (green) and decreased (pink) in relative abundance in the synthetic feed fraction, representing translationally/metabolically active genus. Heatmap sidebars represent square root transformed relative abundances.

Figure 5.8 Fold-changes of active taxa in the synthetic feed compared to the water fed (negative control) fraction (A – E). Point colour indicates taxa that were increased (green) and decreased (pink) in relative abundance in the synthetic feed fraction, representing translationally/metabolically active genus. Heatmap sidebars represent square root transformed relative abundances.

5.3.2 Metaproteomic-based translationally active taxonomic profiling of syntrophic octanoic acid degradation

In this study, metagenomic analysis of BONCAT-ABCS samples reveals the translational activity heterogeneity in AD microbial communities, in which distinct cellular subpopulations can respond differently to the presence of VFAs. The translationally active microbes react towards VFA availability, or other changes in the environment, by constantly adjusting their protein synthesis which will in turn drive their cellular function. To gain insight into the proteins expressed by translationally active cells throughout the course of octanoic acid degradation, proteomic analysis was performed on enriched proteins from BONCAT labelled/unlabelled samples.

BONCAT was used to label nascent proteins with the synthetic amino acid azidohomoalanine (AHA), these proteins were then separated from their unlabelled counterparts via ABCS and subsequently identified and quantified by LC-MS/MS. A total of 14 samples were analysed by LC-MS/MS (Appendix D). Based on metabolic activity shown in Fig. 5.3, BONCAT and non-BONCAT samples 12 h after nutrient rich synthetic feed (positive control) and water fed (negative control) were compared. Spectra were searched against a predicted protein (prokka annotations) database derived from MAGs acquired from the metagenomic data. Once identified, peptides were used to estimate the translationally active taxa at the protein level. Total spectrum count was used as a proxy for translationally active cell abundance.

In total 128 proteins were identified, 49 of which were only present in octanoic acid samples, 37 in synthetic feed samples and 4 in water fed samples (Fig. 5.9). For all proteins identified in the octanoic samples, the total spectrum count abundance was normalised by subtracting the total spectrum count abundance of the relevant non-labelled sample to identify proteins preferentially synthesized in the course of octanoic acid degradation. Complete proteomic results are listed in Appendix H. More than 45% of identified proteins were annotated as 'hypothetical proteins', suggesting that this list contains poorly characterised proteins that play important roles in regulating AD microbes's physiology during octanoic acid degradation or they could be surface expressed proteins.
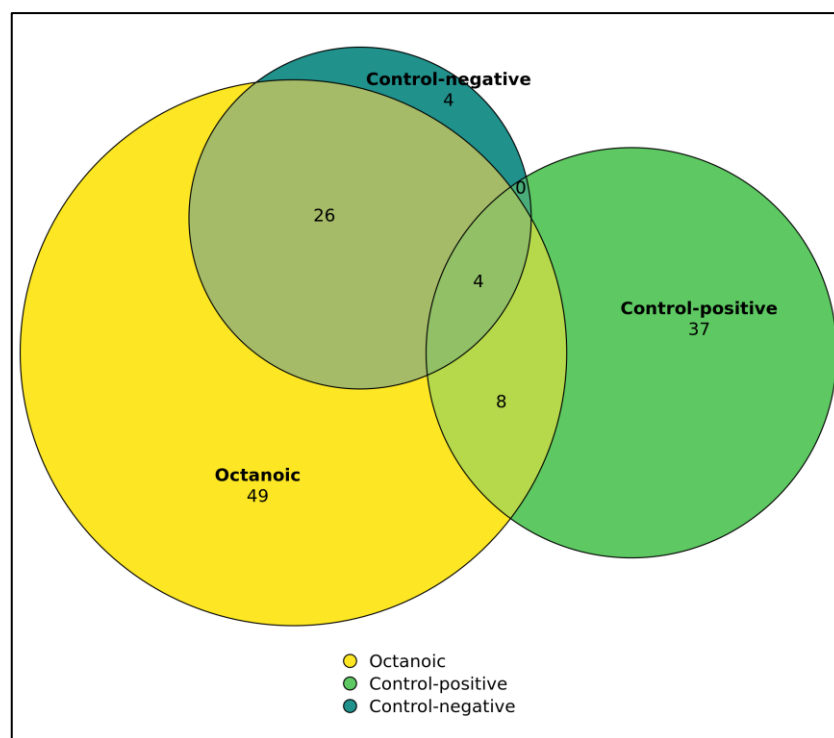
Figure 5.9 Venn diagram showing the unique and common expressed proteins retrieved from BONCAT-ABCS. Low number of unique proteins acquired from water fed (negative control) compared to octanoic acid and synthetic feed (positive control) samples. The complete proteomic results in Appendix H dataset was used to make the Venn diagram.

The acquired prokka accessions were used to estimate the associated taxonomically classified MAGs in order to identify the organisms that synthesised the proteins and were therefore translationally active. The identified genera and abundance in the course of octanoic acid degradation are shown in Fig. 5.10. Before octanoic acid addition, *Methanothrix* (57.6%) and *Methanolinea* (8.3%) were the most abundant and translationally active taxa in VFA-starved communities, followed by *Microbacterium* (5.6%), unclassified *Rickenellaceae* (4.2%), *Clostridium* (3.5%), *Polaromonas* (3.5%) and *Streptomyces* (2.8%). By 6 hours after octanoic acid addition, the enriched labelled proteins were dominated by proteins expressed by 8 genera, such as *Methanothrix* (53.9%), *Methanolinea* (11.2%), *Streptomyces* (8.9%), *Microbacterium* (7.9%), unclassified *Prolixibacteraceae* (3.4%), unclassified *Syntrophaceae* (2.2%) and *Acidovorax* (1.1%). This result indicates the important role of those genera in the early stage of octanoic acid degradation, while *Methanothrix* and *Methanolinea* were the main acetate and hydrogen consumers, respectively. The diversity of translationally active microbes increased two-fold to 15 genera at 12 and 18 hours after octanoic acid addition following the availability of VFA intermediates, such as hexanoic and acetic acid. Proteins expressed by *Polaromonas*, *Clostridium*, *Syntrophorhabdus* and unclassified *Methanomicrobia* were detected during this period of octanoic acid degradation. By 24 hours, the total number of identified spectra decreased, but were still dominated by proteins expressed by methanogens, *Methanothrix* (30.7%) and *Methanolinea* (28.4%), bacterial proteins were dominated by *Streptomyces* (8%), unclassified *Bacteroidales* (6.8%), unclassified *Syntrophaceae* (4.5%), *Tahibacter* (4.5%), *Clostridium* (3.4%) and *Polaromonas* (3.4%).
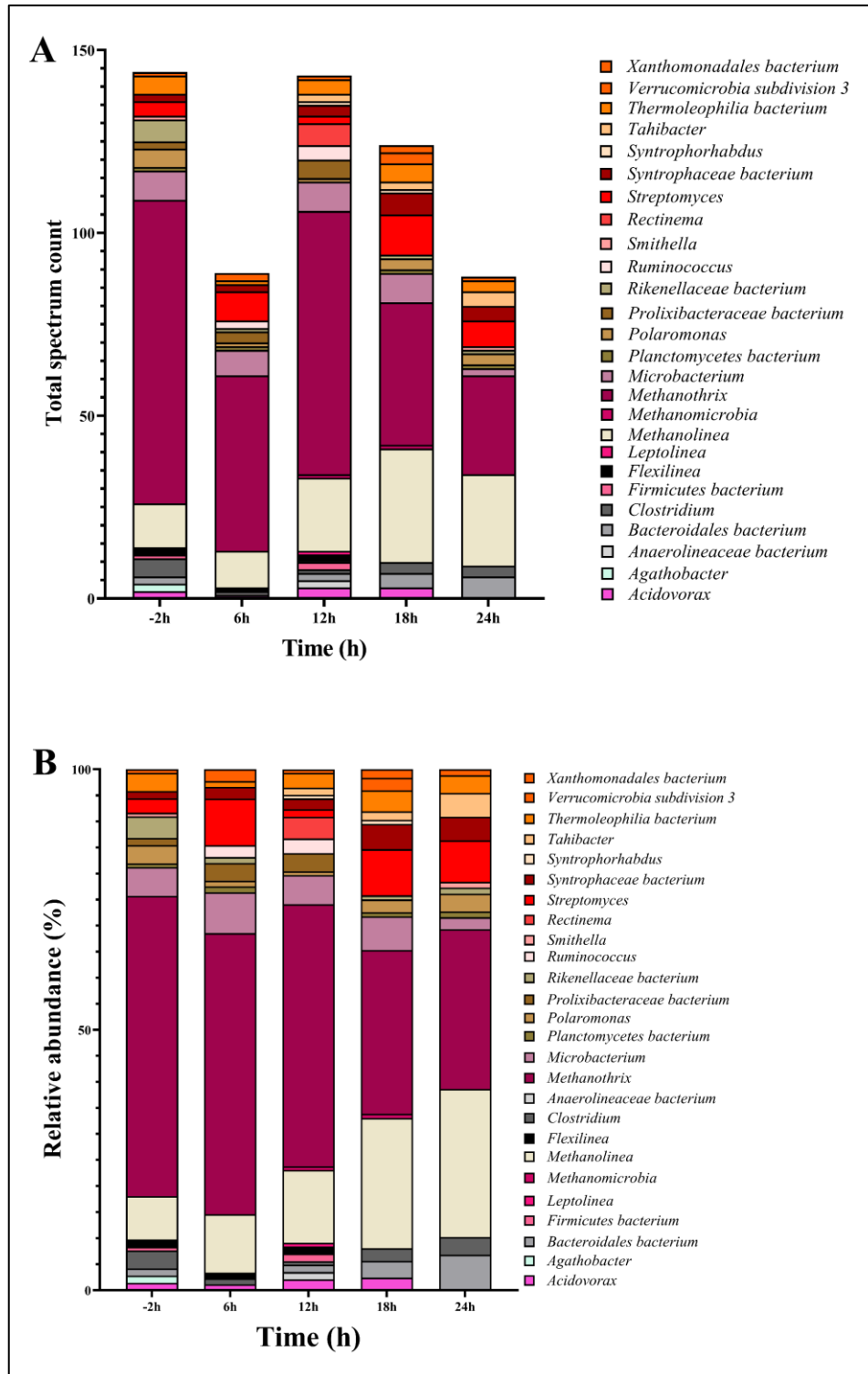
Figure 5.10 Abundance of active genera from proteomic data. The taxa plot showing absolute (A) and relative abundance (B) of identified genera in the course of octanoic acid degradation. The acquired prokka accessions were used to estimate the associated taxonomically classified MAGs in order to identify the translationally active microbes. Total spectrum count used as a proxy of for translationally active cell abundance.

Metaproteomic analysis of labelled and enriched proteins showed differences in community structure based on the substrate added. Methanogens (*Methanothrix* and *Methanolinea)* and *Microbacterium* were the most abundant metabolically active microbes in the octanoic acid fed microbial community, while the nutrient rich fed (positive control) community was dominated by bacteria, such as *Rectinema*, *Ruminococcus* and unclassified members of *Prolixibacteraceeae*, *Anaerolinaceeae* and *Synergistaceae* (Fig. 5.11). Water fed (negative control) AD community was dominated by *Methanothrix* and *Methanolinea,* which was reinforced by their prominence in VFA-starved AD community before octanoic acid addition. This supports their continued metabolic activity in AD systems during starvation.

Proteomic analysis of AHA-labelled and enriched proteins improved the translationally active estimates of our metagenomic analysis. More importantly, it showed that the top three MAGs identified by proteomic-based translationally active taxa profiling (*Microbacterium*, *Methanolinea* and *Methanothrix*) agreed with those revealed by the metagenomic-based profiling (Fig. 5.12). These results also demonstrated that BONCAT-ABCS provided a fair coverage of metabolically active microbes at the proteome level and reflected on the quantity of underrepresented genera that often appear in low abundances in AD microbial communities.
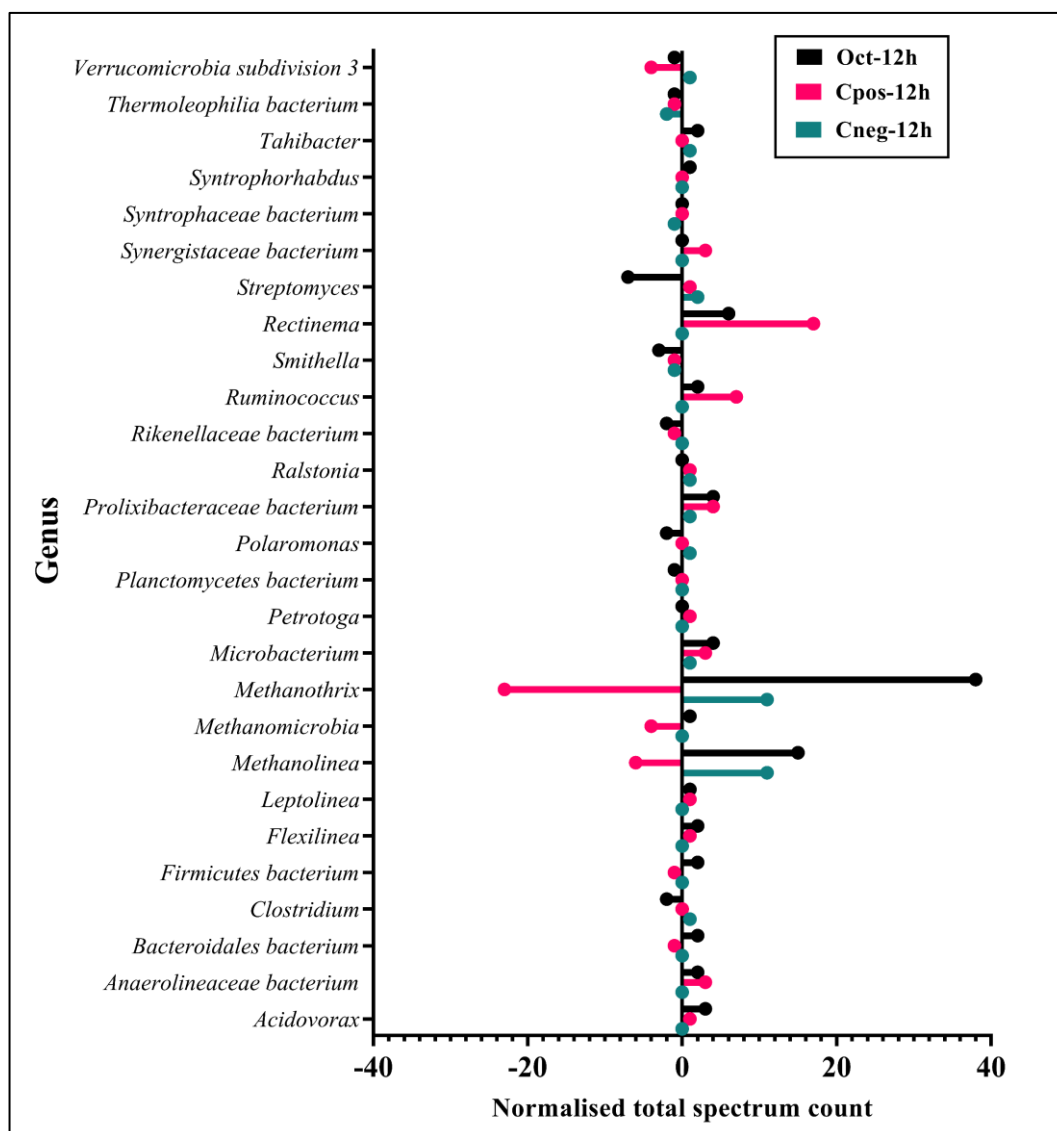
Figure 5.11 BONCAT-ABCS proteomic data revealed various translationally active microbes depending on the substrate added. The labelled and enriched proteins retrieved by BONCAT-ABCS were used to compare the identified genus and their abundance in octanoic, synthetic feed (positive control) and water fed (negative control) bioreactors 12 hours after addition. The total spectrum count abundance was normalised by subtracting the total spectrum count abundance of the corresponding non-labelled sample to indicate taxa translational activity 12 h after substrate addition. Positive numbers indicate increase translational activity, whereas negative indicates low translational activity.
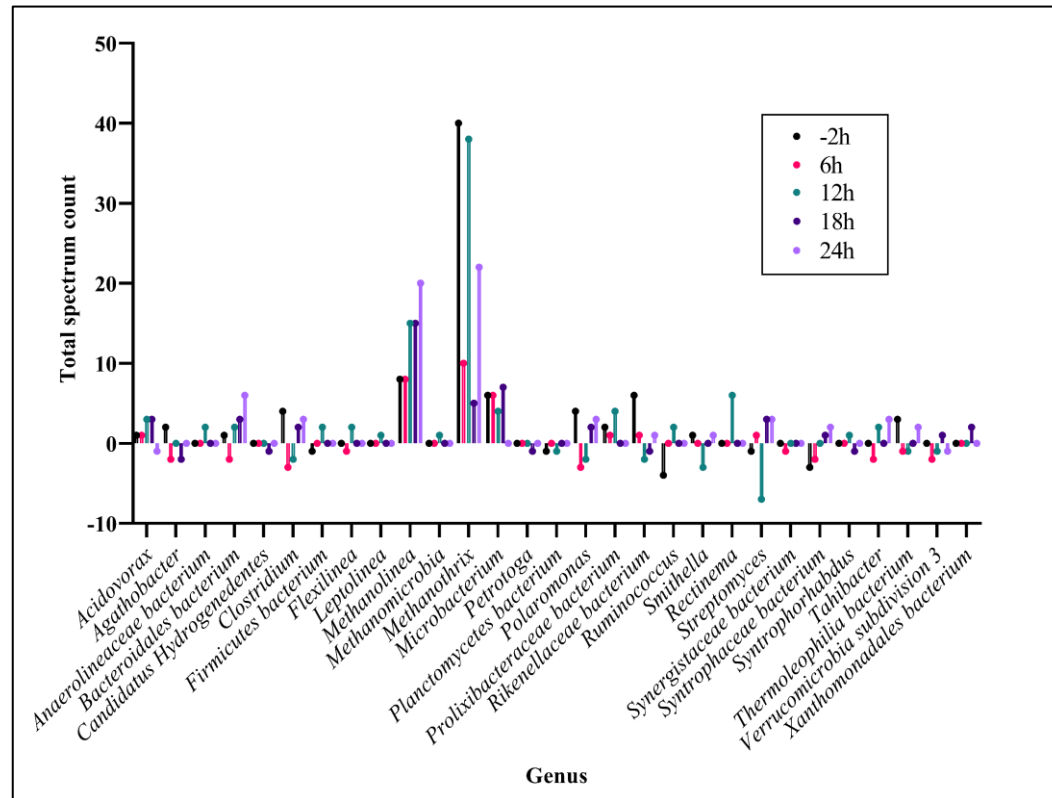
Figure 5.12 A cascade of various organisms become active as substrate become available in the course of octanoic acid degradation. The labelled and enriched proteins retrieved by BONCAT-ABCS were used to compare the identified genus and their abundance in octanoic acid bioreactors. The total spectrum count abundance was normalised by subtracting the total spectrum count abundance of the corresponding non-labelled sample to indicate taxa translational activity during octanoic acid degradation. Positive values imply increased translational activity, whereas negative values suggest decreased translational activity. The colours represent sample points throughout the octanoic acid degradation process.

5.3.3 Metagenomic and proteomic data generated from BONCAT-ABCS reveals the functional landscape of abundant genera in the degradation of octanoic acids

The proteomics analysis of BONCAT-ABCS data allowed the investigation of proteins expressed by each genus, one or a few at a time. Appendix I lists the fractions of proteins in the abundant genera that were detected using metaproteomic data generated in the course of octanoic acid degradation. The Kyoto Encyclopaedia of Genes and Genomes (KEGG; Kanehisa et al., 2016) pathway database was used to identify expressed genes related to fatty acids (pathway map00071) and methanogenesis (pathway map00680) metabolic pathways.

Interestingly, most of the expressed proteins enriched via BONCAT-ABCS were membrane bound proteins, such as ABC transporter substrate-binding proteins (*urtA*) and nitrate reductase (*nar*H). Highly abundant translationally active bacteria detected in VFA-starved AD communities (before substrate addition), such as *Microbacterium* and *Acidovorax*, expressed proteins required for respiratory nitrate reduction (*nar*H), which supported the use of nitrate as a final electron acceptor in anaerobic conditions for the maintenance of a proton motive gradient to continue growing (Sohaskey and Wayne, 2003). While others, such as *Clostridium*, *Streptomyces*, *Polaromonas* and unclassified *Rickenellaceae*, expressed elongation factor thermal unstable Tu (*tuf*) protein, which catalyses the binding of aminoacyl-tRNA to the A-site of the ribosome inside living cells and has evolved the capacity to execute diverse functions on the extracellular surface of prokaryotic cells (Sprinzl, 1994; Harvey et al., 2019), and bacterial transport proteins, including ATP-binding cassette (ABC) transporters (*urtA*) to import and expel substrates.

Microbacterium was identified as the most abundant bacteria based on metagenomic and proteomic data from octanoic acid BONCAT-ABCS samples (Fig 5.7 and 5.12). Genes encoding for the entire fatty acid β-oxidation cycle were identified within *Microbacterium* published genome (Fig. 5.13) suggesting its important role in octanoic acid degradation. Genes for acyl-CoA oxidase (ACO) (EC 1.3.3.6) which catalyses the Cα-Cβ oxidation of fatty acids and is active on CoA derivatives of fatty acids with aliphatic chains from 8 to 18 carbons (Martin et al., 2020). Enoyl-CoA hydratase (ECH) (EC:4.2.1.17) catalyses the second step in β-oxidation pathway, the hydration of the bond between C-2 and C-3, resulting in the formation of a β-hydroxyacyl-CoA thioester (Agnihotri and Liu, 2003). Acetyl-CoA acyltransferase (EC:2.3.1.16) (EC:2.3.1.9) catalyses the final step of fatty acid oxidation in which acetyl-CoA is released and the CoA ester of a fatty acid two carbons shorter is formed. *Microbacterium* is able to grow anaerobically using nitrate as a terminal electron acceptor as it also expressed genes encoding nitrate reductase (EC:1.7.5.1 1.7.99.-) involved in the nitrogen metabolism (map00910) in the course of octanoic acid degradation.
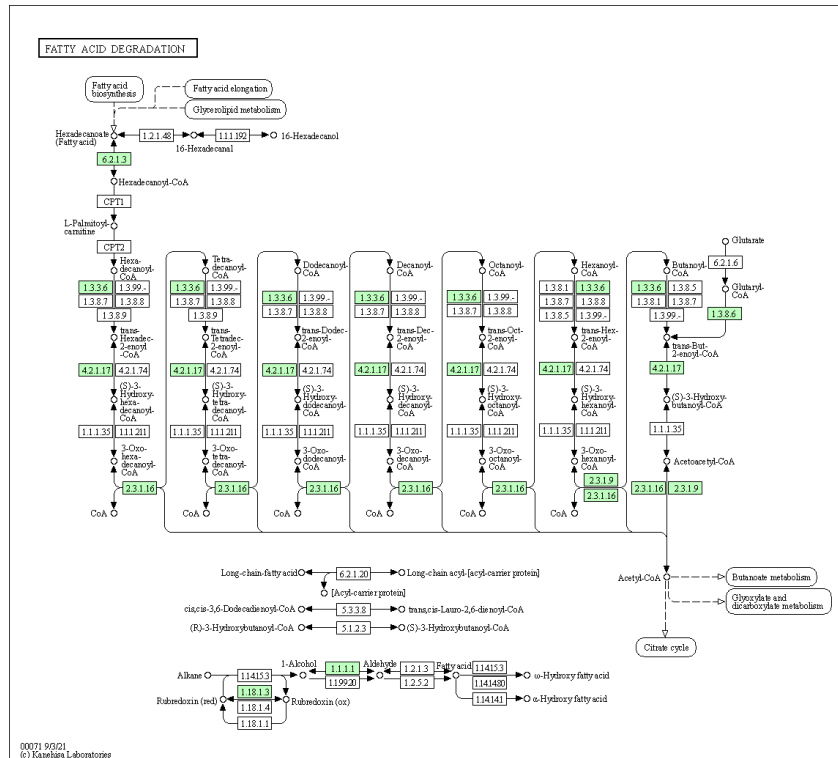
Figure 5.13 *Microbacterium* fatty acid degradation pathway. The β-oxidation pathway in the KEGG pathway map (map00071), with the detected genes in the *M. testaceum* genome are marked in green (Kanehisa et al., 2016).

Streptomyces showed an increase in abundance in the course of octanoic acid degradation (Fig. 5.12). *Streptomyces* is an obligate aerobic actinobacterium, but could remain metabolically active under anaerobic conditions by synthesising respiratory nitrate reductase (*Nar*) enzymes that contribute to the maintenance of membrane potential and energy conservation (Sawers et al., 2019). *Streptomyces* is able to oxidise long- and medium-chain fatty acids as it published genome contains genes encoding acyl-CoA oxidase (EC:1.3.3.6) and acyl-CoA dehydrogenase (EC:1.3.8.7), respectively (Fig. 5.14). Moreover, it also has the genes that encode enzymes that catalyse the hydration, NAD^+ oxidation and thiolysis steps in fatty acid β-oxidation, such as enoyl-CoA hydratase (EC:4.2.1.17), 3-hydroxyacyl-CoA

dehydrogenase (EC:1.1.1.35) and acetyl-CoA acyltransferase (EC:2.3.1.16) (EC:2.3.1.9), respectively (Fig. 5.14).
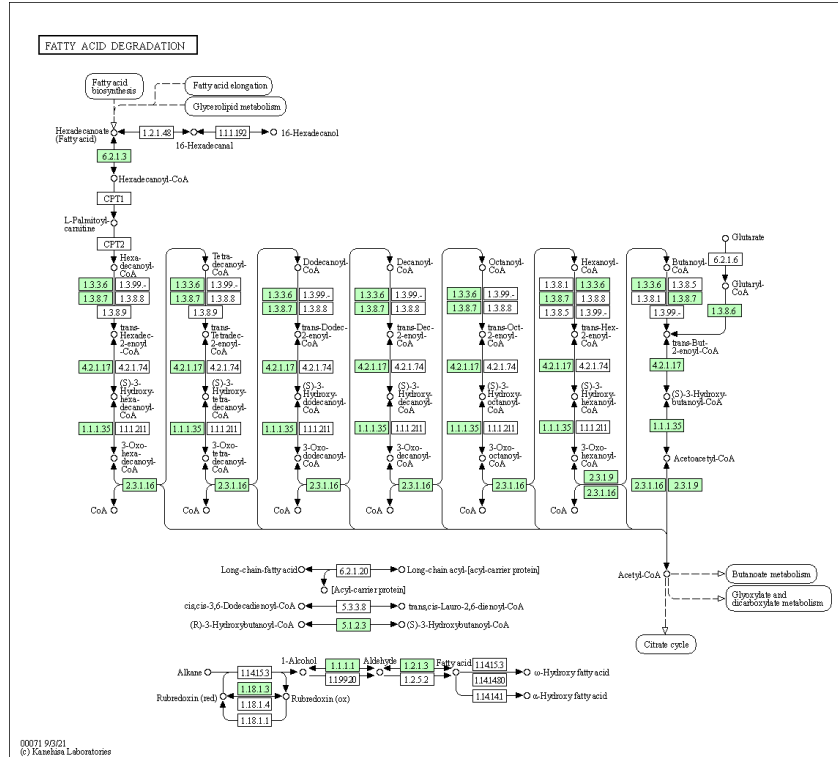
Figure 5.14 *Streptomyces* fatty acid degradation pathway. The β-oxidation pathway in the KEGG pathway map (map00071), with the detected genes in the *S. coelicolor* genome are marked in green (Kanehisa et al., 2016).

Acidovorax, is a genus of Proteobacteria. All species are facultative anaerobes and grow anaerobically using nitrate as a terminal electron acceptor. This is confirmed by the presence of proteins required for respiratory nitrate reduction (*narH*) in the proteins fraction enriched by BONCAT-ABCS. Genes encoding the entire fatty acid β-oxidation cycle were identified within *Acidovorax* published genome (Fig. 5.15) suggesting its important role in octanoic acid degradation. *Acidovorax* shows increase labelled proteins in the course of octanoic acid degradation (Fig. 5.12). *Acidovorax* genomes contains genes encoding medium-chain acyl-CoA dehydrogenase (EC:1.3.8.7) and short-chain acyl-CoA

dehydrogenase or butyryl-CoA dehydrogenase (EC:1.3.8.1) suggesting its ability to metabolise medium- and short-chain fatty acids by catalysing the formation of a double bond between the Cα and Cβ. Enoyl-CoA hydratase (EC:4.2.1.17) catalyses the second step in β-oxidation pathway and 3-hydroxyacyl-CoA dehydrogenase (EC:1.1.1.35) catalyses the third step of β-oxidation, converting the hydroxyl group into a keto group. Acetyl-CoA acyltransferase (EC:2.3.1.16) (EC:2.3.1.9) catalyses the final step of β-oxidation.
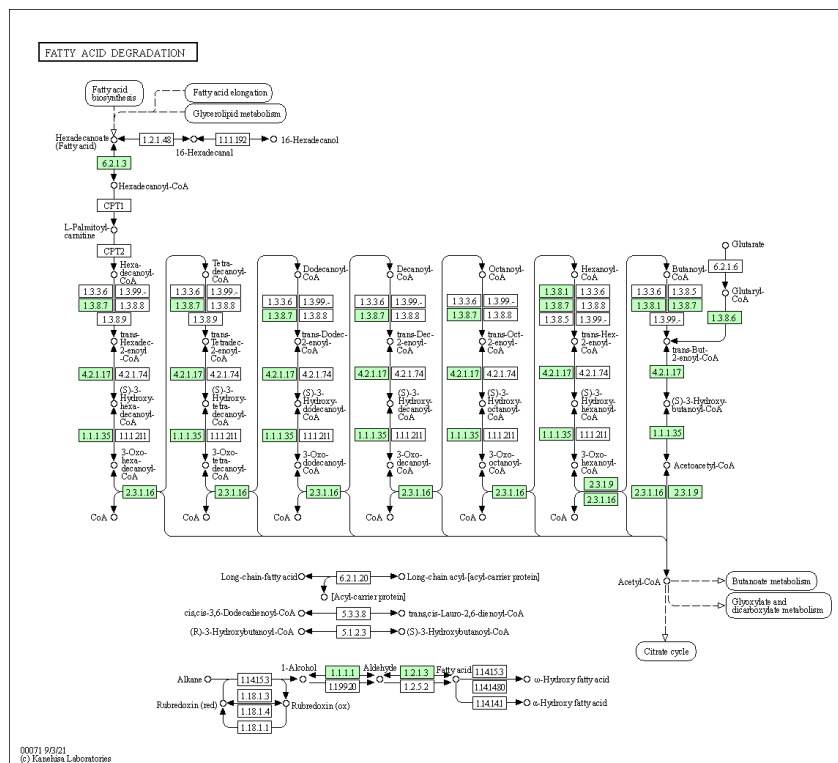
Figure 5.15 *Acidovorax* fatty acid degradation pathway. The β-oxidation pathway in the KEGG pathway map (map00071), with the detected genes in the *A. citrulli* genome are marked in green (Kanehisa et al., 2016).

Polaromonas is a chemoorganotrophic proteobacterium that could oxidise medium- and small chain fatty acids (Fig. 5.16). *Polaromonas* becomes more active during 18 - 24 h of octanoic acid degradation (Fig. 5.12). Its published genome contains genes encoding acyl-CoA dehydrogenase (EC:1.3.8.7), butyryl-CoA

dehydrogenase (EC:1.3.8.1), enoyl-CoA hydratase (EC:4.2.1.17), 3-hydroxyacyl-CoA dehydrogenase (EC:1.1.1.35), acetyl-CoA acyltransferase (EC:2.3.1.16) and acetyl-CoA C-acetyltransferase (EC:2.3.1.9).
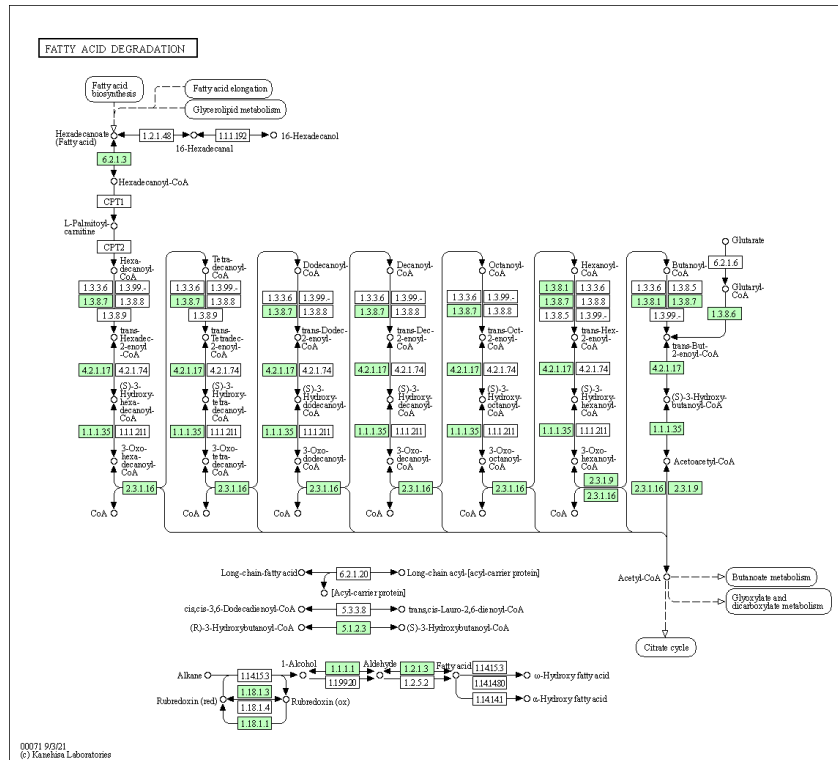
Figure 5.16 *Polaromonas* fatty acid degradation pathway. The β-oxidation pathway in the KEGG pathway map (map00071), with the detected genes in the *P. naphtalenivorans* genome are marked in green (Kanehisa et al., 2016).

Clostridium are anaerobic, fermentative, spore-forming Gram-positive bacteria belonging to the phylum Firmicutes. *Clostridium* becomes increasingly active as octanoic acid breakdown progresses (18 – 24 hours; Fig. 5.12). *Clostridium* are only able to metabolise short-chain fatty acids as they only possess the genes encoding short-chain acyl-CoA dehydrogenase or butyryl-CoA dehydrogenase (EC:1.3.8.1) and acetyl-CoA C-acetyltransferase (EC:2.3.1.9) in their genomes (Fig. 5.17). Butyric acid is a 'genus specific' product of fermentation.
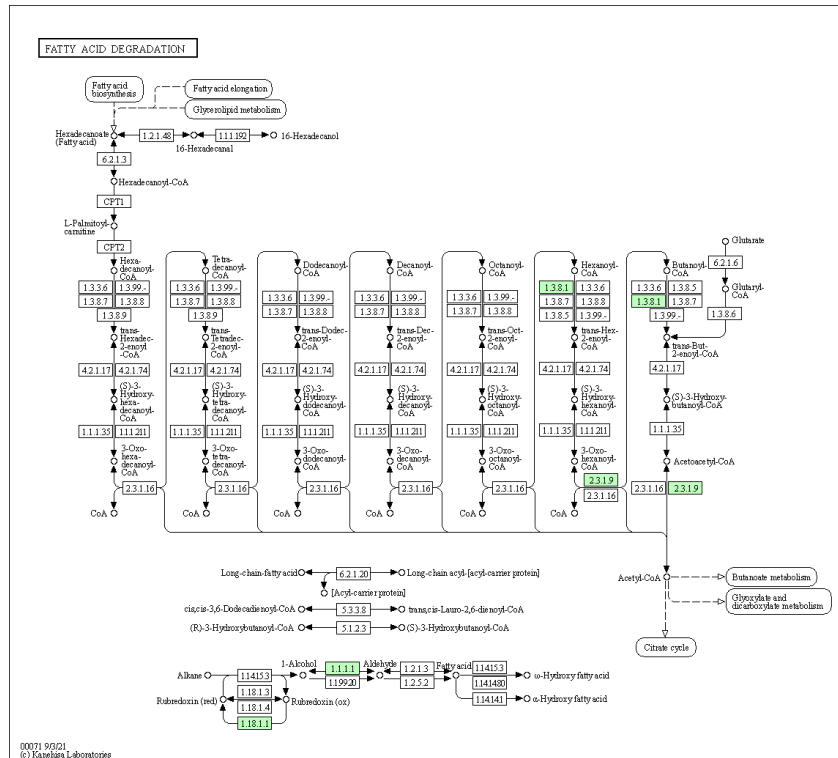
Figure 5.17 *Clostridium* short-chain fatty acid degradation pathway. The KEGG pathway map (map00071) for the β-oxidation, with the identified genes in the *C. acetobutylicum* genome highlighted in green (Kanehisa et al., 2016).

Syntrophic conversion of octanoic acid under anaerobic conditions involves interspecies electron transfer between bacteria and archaea, which commonly involves hydrogen and/or formate as electron shuttles (Ziels et al., 2015). The formate dehydrogenase major subunit (EC:1.17.1.9) was detected in the enriched proteins expressed by *Methanolinea* (Fig. 5.18). Methyl-coenzyme M reductase alpha/beta/gamma subunits (EC:2.8.4.1), which are responsible for catalysing the terminal step in methanogenesis, and tetrahydromethanopterin S-methyltransferase subunit A (EC:2.1.1.86), which has a key function in energy conservation by catalysing the methyl transfer from methyl-tetrahydromethanopterin to coenzyme M and its coupling with sodium-ion translocation, were also expressed by *Methanolinea* during octanoic acid degradation. This suggests the

hydrogenotrophic methanogen, *Methanolinea*, as one of the main hydrogen and/or formate consumers in the course of octanoic acid degradation (Fig. 5.12).
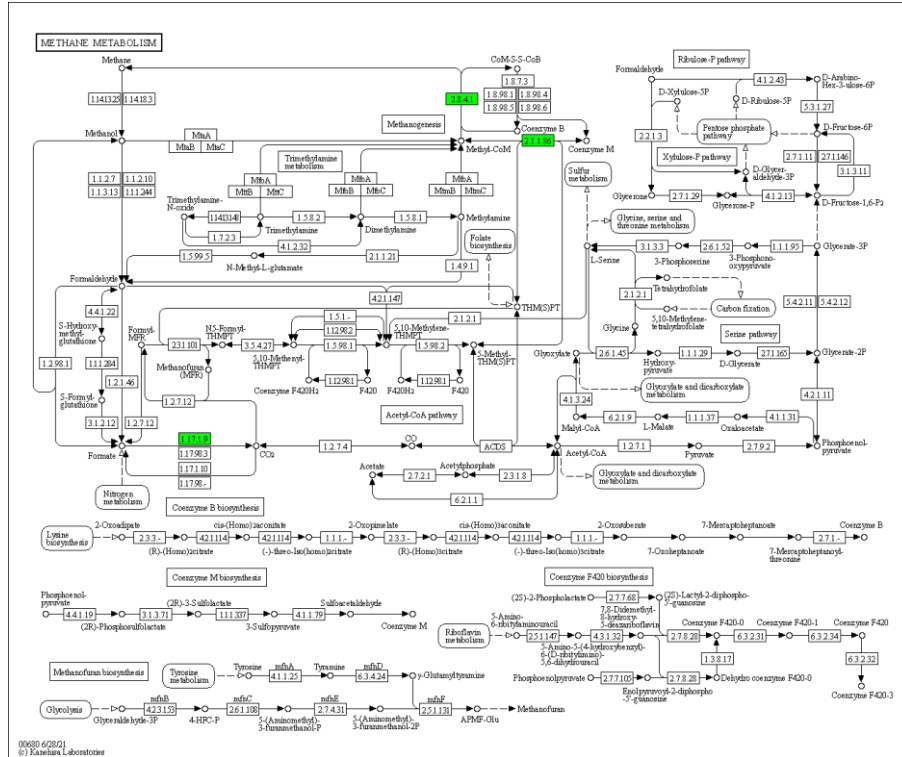
Figure 5.18 *Methanolinea* methane metabolism pathway. The KEGG pathway map (map00680) shows the methane metabolism reference pathway, with *Methanolinea* expressed proteins highlighted in green (Kanehisa et al., 2016).

Several *Methanothrix* expressed proteins involved in acetoclastic methanogenesis (Fig. 5.19) were detected, which shows its important role as an acetate consumer in the syntrophic oxidation of octanoic acid. Methyl-coenzyme M reductase alpha/beta/gamma subunits (EC:2.8.4.1), acetyl-CoA decarbonylase/synthase, CODH/ACS complex subunit beta (EC:2.3.1.169) and acetyl-CoA synthetase (EC:6.2.1.1) were detected in the BONCAT-ABCS enriched proteins and are central to the acetate switch and catalyse the overall process of acetate production or acetate utilization. BONCAT-ABCS also detected the anaerobic carbon-monoxide dehydrogenase, CODH/ACS complex subunit alpha

(EC:1.2.7.4), which alongside acetyl-CoA decarbonylase/synthase, CODH/ACS complex subunit beta (EC:2.3.1.169) catalyses the acetyl coenzyme A (CoA) pathway, commonly referred to as the Wood-Ljungdahl pathway. This specific pathway is characterised by the use of hydrogen as an electron donor and carbon dioxide as an electron acceptor to produce acetyl-CoA as the final product. The carbon monoxide dehydrogenase allows *Methanothrix* to use carbon dioxide as a source of carbon and carbon monoxide as a source of energy. It suggests that *Methanothrix* is able to utilize the acetyl-CoA pathway to fix carbon dioxide. Additionally, the V/A-type H^+/Na^+-transporting ATPase subunit A [EC:7.1.2.2 7.2.2.1], which is a membrane-bound ATP synthase used to synthesise ATP via a proton or sodium ion gradient, was also detected.
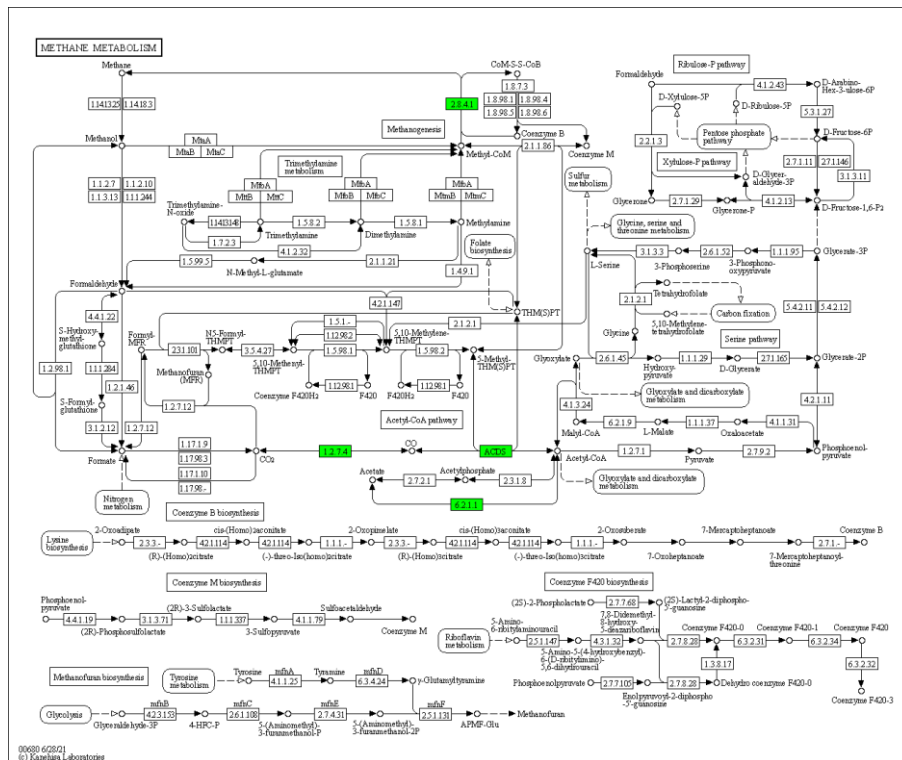
Figure 5.19 *Methanothrix* methane metabolism pathway. The KEGG pathway map (map00680) shows the methane metabolism reference pathway, with *Methanothrix* expressed proteins highlighted in green (Kanehisa et al., 2016).

5.4 Conclusion

BONCAT can be used to track the active members of an AD community in response to changes in metabolic conditions. By using affinity-based click chemistry, active microbes were selectively recovered for DNA and protein extraction. The analysis of metagenomic and proteomic sequences from the AHA-based BONCAT-ABCS samples confirmed that this method enriched distinct microbial populations from the total community, allowing us to begin to infer which microbes in the AD community samples were most likely involved in octanoic acid degradation. Targeted analysis of a portion of these data confirmed that the method did not just recover the most abundant species within the community but was selective for organisms with an important role in the metabolism of available substrates. Further utilisation of BONCAT-ABCS as a rapid tool for capturing changes within complex microbial systems should enable us to further our understanding of the metabolic community interactions occurring within anaerobic digestion.

6 Discussion and future work

6.1 Discussion

The anaerobic oxidation of even- and odd-chain medium length fatty acids by VFA-starved AD communities generates a variety of intermediate metabolites, which may influence the activity of syntrophic bacteria and methanogenic archaea whose presence was detected by BONCAT. Expanding on recent studies that have employed BONCAT to characterise the ecophysiology of microbial communities in their natural growth environment (Hatzenpichler et al., 2014; Samo et al., 2014; Hatzenpichler et al., 2016; Leizeaga et al., 2017; Couradeau et al., 2019; Valentini et al., 2020), this approach was used in combination with ABCS, metagenomic and proteomic-based approaches to unravel the *in situ* metabolic activities in an AD community during the degradation of medium-chain fatty acids and enhance our understanding of microbial community interactions in AD systems.

During the diauxic growth of mixed *E. coli* strains, the BONCAT-ABCS approach was successfully applied and displayed the specificity and sensitivity to enrich active cells. The incorporation of biotin-alkyne moieties into some cells and the selective enrichment of these onto neutravidin beads is consistent with the labelling reaction resulting in a fraction of nascent proteins locating to the surface of metabolically active *E. coli* cells. The effective enrichment of AHA-based BONCAT followed by ABCS suggests that this is a feasible method for isolating populations of microorganisms in mixed systems that are metabolically responsive to substrate availability.

The BONCAT-ABCS metagenomic data reveal the extensive heterogeneity of translational/metabolic activity among members of AD microbial communities. Each treatment (octanoic acid, synthetic feed and water fed) harboured a unique microbial community consisting of bacteria and archaea. BONCAT-ABCS sequencing data show that the most abundant bacteria, such as *Microbacterium* and *Streptomyces*, are also active in situ, suggesting an important role for these species in octanoic acid degradation. However, this trend of high relative abundance with high activity did not always hold true. For example, unclassified *Clostridiales* (Firm-04), was highly abundant in all octanoic samples, but had lower relative translational activity than the AD microbial community it co-colonised. This was also found in the octanoic and synthetic feed fractions for unclassified *Bacteroidales* and *Verrucomicrobia subdivision 3*, which were among the most abundant taxa but had low translational activity. Moreover, BONCAT-ABCS also discovered that low abundance bacteria (<1%) were active, indicating that activity in these less abundant species may also be important in the course of octanoic acid degradation. *Syntrophomonas*, for example, exhibited an average relative abundance of 0.1% during the octanoic acid degradation, but showed high translational activity, being 2–4 times more prevalent in octanoic samples than the negative control. Thus, community presence and abundance are not always indicative of translational activity. The enriched fraction isolated on neutravidin beads might also be a better proxy for the intact cellular fraction of the AD microbiome as this method filters out extracellular DNA, necromass and other cells that are not expressing azide-containing proteins on their cell surfaces but might still be found in sequencing results of DNA extracted from AD sludge.

The proteomic workflow for identifying BONCAT-ABCS enriched protein fragments enables us to find the best match to predicted proteins acquired from metagenomic data. This help us to further link the genomic identities of translationally/metabolically active microbes to their functions in AD microbial communities during octanoic acid degradation. The KEGG pathway map (Kanehisa et al., 2016) serve as useful sources of prior knowledge about the fatty acid degradation or methane metabolism pathway of identified taxa inferred from published genomes. It is worth noting that the inferred example species probably not be completely representative of the organisms in an AD community.

Based on the observed activity and the presence of genes involved in fatty acid breakdown or methane metabolism as extrapolated from their inferred published genomes, a cascade of various organisms may become active as substrate become available. *Microbacterium* and *Acidovorax* seemingly played important syntrophic role at the beginning of octanoic acid degradation (Fig 5.12). *M. testaceum* published genome encoding gene for acyl-CoA oxidase (ACO) (EC 1.3.3.6) which catalyses the Cα-Cβ oxidation of fatty acids and is active on CoA derivatives of fatty acids with aliphatic chains from 8 to 18 carbons (Martin et al., 2020). *M. testaceum* lack the gene encode 3-hydroxyacyl-CoA dehydrogenase (EC:1.1.1.35), which is involved in the oxidation of L-β-hydroxyacyl CoA by NAD^+, while *Acidovorax* has it in their published genome (*e.g., A. citrulli*). *Microbacterium* and *Acidovorax* activity release VFA intermediates, such as hexanoic, butyric and acetic acid (Fig. 5.3). The presence of those VFA intermediates seemingly induces the activity of other microbes, such as *Streptomyces*, *Polaromonas* and *Clostridium* (Fig. 5.12). A complete pathway for medium- and short-chain fatty acid β-oxidation

was found in *Streptomyces* (*e.g. S. coelicolor*) and *Polaromonas* (*e.g. P. naphthalenivorans*), indicating their capability of metabolising the intermediates fatty acids produced in the course of octanoic acid degradation. *Clostridium*, which has the complete short-chain fatty acid oxidation pathway in its published genome (*e.g. C. acetobutylicum*), was also enriched via BONCAT-ABCS in the course of octanoic acid degradation. It is strongly suggesting that *Clostridium* population was enriched in AHA-based BONCAT labelled protein due to cross-feeding of shorter-chain intermediates (*e.g.* butyric acid) produced during octanoic-acid β-oxidation.

To complete syntrophic octanoic oxidation, NAD^+ and FAD^+ must be regenerated. The supply of electron carriers (FAD^+ and NAD^+) are regenerated by either a soluble hydrogenase that catalyses the oxidation of highly reduced compounds and releases molecular hydrogen (H_2) or by a cytoplasmic formate dehydrogenase (FDH) that act as CO_2 reductase to form formic acid (Agne et al., 2021). Hydrogen and formic acid produced during β-oxidation are consumed by hydrogenotrophic methanogens to form methane. BONCAT-ABCS was able to enrich proteins closely matched to predicted *Methanolinea* proteins. A genomic analysis of the *Methanolinea* MAGs indicated that its harboured gene encoding enzymes for methane production from hydrogen and/or formic acid. This observation agrees with the physiology of *Methanolinea*, which are known as hydrogenotrophic methanogens (Imachi et al., 2008). Genes encoding formate dehydrogenase major subunit (EC:1.17.1.9), methyl-coenzyme M reductase alpha/beta/gamma subunit (EC:2.8.4.1) and tetrahydromethanopterin S-methyltransferase subunit A (EC:2.1.1.86) were present in the MAGs most closely matched to *Methanolinea*.

The most closely matched BONCAT-ABCS enriched protein fragments to predicted *Methanothrix* proteins suggest that this species plays important role in the synthesis of methane from acetate (Huser et al., 1982) in syntrophic octanoic acid oxidation. *Methanothrix* MAGs contained acetyl-CoA synthetase (EC:6.2.1.1) for acetate activation, acetyl-CoA decarbonylase/synthase, CODH/ACS complex subunit beta (EC:2.3.1.169) to oxidatively split acetyl-CoA into CO_2 and methyl-coenzyme M reductase alpha/beta/gamma subunit (EC:2.8.4.1) for methyl-CoM reduction to methane. The BONCAT-ABCS also enriched the anaerobic carbon-monoxide dehydrogenase, CODH/ACS complex subunit alpha (EC:1.2.7.4), which alongside acetyl-CoA decarbonylase/synthase, CODH/ACS complex subunit beta (EC:2.3.1.169) catalyse the acetyl coenzyme A (CoA) pathway, commonly referred to as the Wood-Ljungdahl pathway. This suggests that *Methanothrix* is able to utilize the acetyl-CoA pathway to fix carbon dioxide, which is at odds with the notion that this species is an obligate acetoclastic methanogens (Huser et al., 1982). The existence and expression of the CO_2-reducing pathway in *Methanothrix* have been previously described (Zhu et al., 2012; Rotaru et al., 2014a; Holmes et al., 2017), and it was hypothesised to be involved in DIET-mediated methane formation. However, the mechanism through which *Methanothrix* directly accepts electrons from its syntrophic partner has not been identified (Rotaru et al., 2014; Holmes et al., 2017; Ziels et al., 2019).

The BONCAT-ABCS also revealed that the enriched proteins fragments closely matched to predicted *Microbacterium*, *Acidovorax*, *Methanolinea* and *Methanothrix* were abundant in VFA-starved AD community before octanoic acid addition. *Microbacterium* and *Acidovorax* seemingly adopting a persister-like

strategy in which reduced cellular activity confers an anoxic-resistant phenotype, allowing them to survive anaerobic conditions and persist during starvation (Sohaskey and Wayne, 2003). *Methanolinea* has been shown to thrive in conditions with low H_2 evolution rates (Imachi et al., 2008). *Methanothrix* has greater substrate affinity and therefore is prevalent in digesters with low acetate (<1 mM) availability (Gujer and Zehnder, 1983; Conklin et al., 2006; Karakashev et al., 2005; Razaviarani and Buchanan, 2014). Moreover, while active cells are more likely to be involved in octanoic acid degradation, inactive cells are equally important because bacteria do not need to be translationally/metabolically active to have an impact on their surrounding community. It is known, for example, that mostly inactive populations can drive geochemical process in their environment (Hausmann et al., 2019; Valentini et al., 2020). Nutrient exchange and production of virulence agents and tiny metabolites are all ways whereby translationally inactive cells might shape their growing environment (Valentini et al., 2020). Microbial necromass (non-living microbial biomass) in an AD community could play an important role in providing a significant source of carbon and nitrogen (Dong et al., 2021). Further characterisation of activity heterogeneity and the contributions of both active and inactive populations in AD system is important to understand the full complexity of the AD microbiome and its available competitive and syntrophic metabolisms, to optimise energy recovery from AD.

Though BONCAT is a great tool to study the AD microbial community, it does have several limitations, some of which have already been mentioned (Hatzenpichler et al., 2014; Hatzenpichler et al., 2016). First, the affinity-based cell sorting approach to enriching whole cells is imperfect, as cell sorting was done

directly on AD biomass and not on individualised cells. It is possible that inactive cells stick to active cells in the enriched fractions, consistent with the formation of biofilm-based cell aggregations. A biofilm-based cell aggregation in AD is formed by biotic and abiotic interactions between microorganisms and sludge particles, leading to the formation of very compact spherical-shaped aggregates with a diameter of 1 – 3 mm (Wilén et al., 2018) where the microbial cells are self-immobilized in a matrix of extracellular polymeric substances. However, it is non-trivial to separate cells from other materials due to the nature of AD samples, which are compact and contained densely packed cells. Secondly, because the most crucial limitation of BONCAT is its reliance on uptake mechanisms, it cannot be ruled out that some AD microbes seemingly did not take up AHA into their cells or proteins due to the lack of appropriate transporters, the high selectivity of their MetRS, or a high Met/AHA ration in the cytoplasm (Hatzenpichler et al., 2014).

Despite these limitations, BONCAT can be used to extend our understanding of the role of specific microbiota in an AD microbial community. These findings demonstrated that BONCAT is a powerful tool for the visualization and identification of translationally/metabolically active microbes in anaerobic digestion system. Metagenomic and proteomic data generated from AHA-based BONCAT-ABCS could help to resolve the functional landscape by obtaining confirmatory assimilation of azidohomoalanine (AHA) into the nascent proteins of metabolically responsive microbes in an AD community. These findings provide a framework to further link the genomic identities of anaerobic bacteria and archaea with their function within AD microbial community that drive anaerobic degradation of organic waste to yield biogas.

6.2 Future work

The application of BONCAT-ABCS to identify active microbes in the course of odd-chain medium length fatty acids (heptanoic or nonanoic acid) would be interesting, as initial data (Fig. 3.8 and 3.10) shows different VFA profiles and kinetics compared to even-chain fatty acid (octanoic acid). It may show variations in the active microbes involved and different genes are required in the OCFAs β-oxidation. Different concentration (high ≥ 10 mM and low ≤ 5 mM) could be tested in this experiment including the use of synthetic feed and water as positive and negative control, respectively.

Data in Chapter 3 shows that the propionic acid (C_3) decomposition rate was slow and appeared to be the rate limiting step in OCFAs degradation rate. Further experiments to differentiate whether increased methanogen numbers or induction of other enzymes required for C_3 oxidation is essential to increase C_3 degradation rate.

It is also important to explore whether cell detachment prior to the affinity-based cell separation (ABCS) step would show increased selectivity of BONCAT-ABCS of labelled intact cells for metagenomic analysis. BONCAT-labelled cells detachment could remove single or small clusters of labelled cells from the sludge particles. It allows click-labelling and ABCS to be performed on single or small clusters of cells and reduce the possibility of unlabelled cells to be found in sequencing results of DNA extracted from AD sludge.

In summary, bioorthogonal non-canonical amino acid tagging (BONCAT) and affinity-based cell separation (ABCS) was used to identify subsets of metabolically

specialised microbes as they respond to substrate availability in AD systems. The results demonstrate the specific labelling, visualisation and separation of microbes that actively participate in volatile fatty acid (VFA) degradation and suggest a differential response to octanoic acid by specialists within the microbial community. Analysis of metagenomic sequences from a time series of BONCAT-ABCS samples reveals that this method enriched a distinct microbial community with genetically-derived metabolisms consistent with and changing according to the observed metabolic outcomes. Proteomic data generated from AHA-based BONCAT-ABCS help to resolve the functional landscape of enriched microbial community and give better resolution of identified translationally/metabolically active taxa. This enrichment approach allows us to determine the temporal response of those microbes most likely to engage in octanoic acid degradation in AD sludge. This method can be applied to the identification of specialist microbes with a role in degradation of a range of other compounds in AD, enhancing our understanding of microbial community interactions and facilitating the development of strategies for process optimisation.

Appendix A. Cell surface labelling and enrichment of active cells in AD sample via BONCAT-ABCS. Created using Biorender.com.
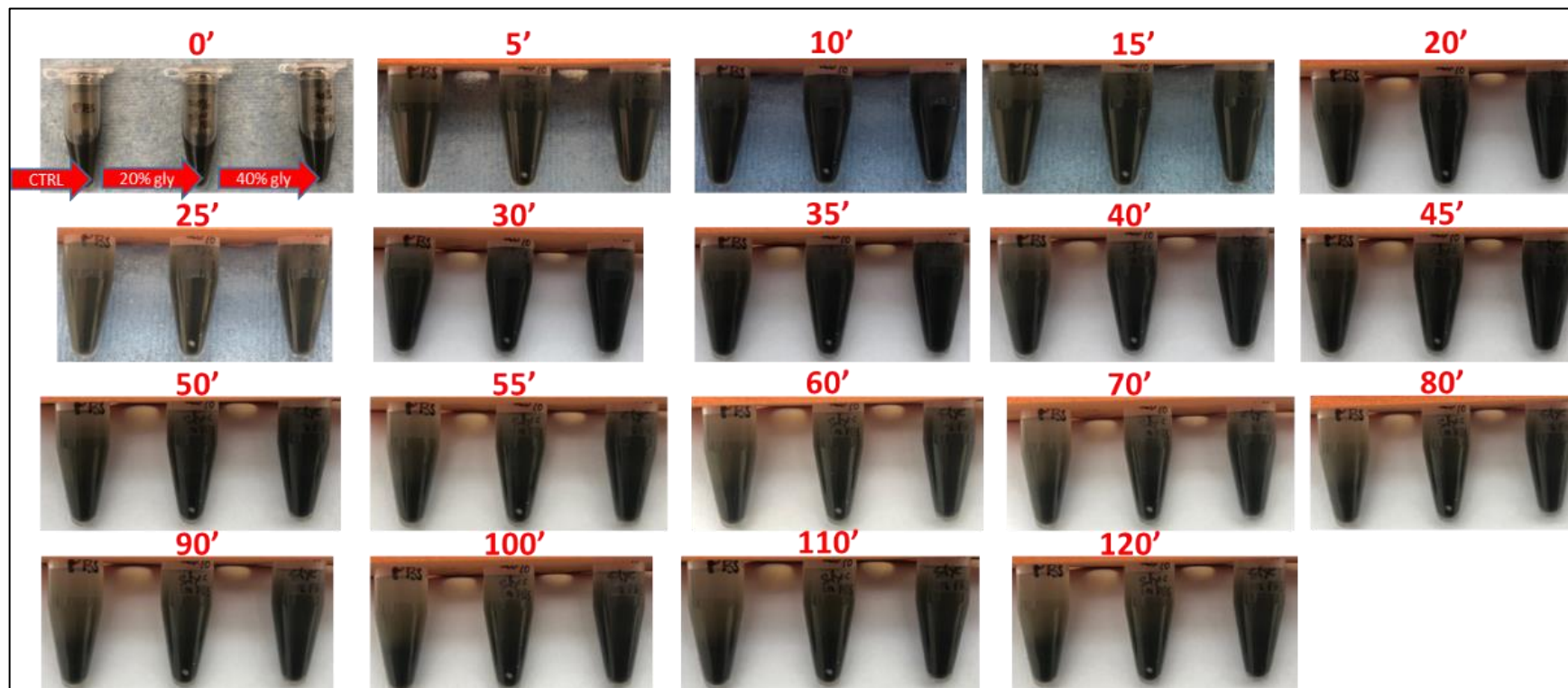
Appendix B. Increase sample density using glycerol to improve magnetic separation. 200 µL of AD sample added with 800 µL PBS or 20% glycerol in PBS or 40% glycerol in PBS. Vortexed for 5 seconds and allowed to stand for 2 hours. Observation was carried out every 5 minute.

Appendix C

Table 1. DNA concentration of samples sent for sequencing using Nanopore and Illumina platforms.

Substrate	Sample (h)	Grouping	dsDNA concentration (ng/µL)	Total amount of DNA (ng)
	-2	Non-BONCAT*	24.3.8	2380
	6	Non-BONCAT*	28.4	2840
	12	Non-BONCAT*	29.8	2980
	18	Non-BONCAT*	31.8	3180
	24	Non-BONCAT*	30	3000
	-2	BONCAT**	4.3	215
	6-R1	BONCAT**	1.37	137
	6-R2	BONCAT**	1.4	140
	6-R3	BONCAT**	1.76	176
Octanoic acid (C₈)	12-R1	BONCAT**	1.377	179.01
	12-R2	BONCAT**	3.98	199
	12-R3	BONCAT**	1.78	178
	18-R1	BONCAT**	1.16	116
	18-R2	BONCAT**	1.31	131
	18-R3	BONCAT**	1.05	105
	24-R1	BONCAT**	1.3	130
	24-R2	BONCAT**	1.53	153
	24-R3	BONCAT**	1.79	179
	-2	Non-BONCAT*	40	4000
Synthetic feed	6	Non-BONCAT*	18.1	1810
(positive control)	12	Non-BONCAT*	37.8	3780
	18	Non-BONCAT*	32	3200
	24	Non-BONCAT*	34.6	3460
	-2	BONCAT**	1.55	155
	6-R1	BONCAT**	2.86	286
	6-R2	BONCAT**	2.1	210
	6-R3	BONCAT**	1.54	154
	12-R1	BONCAT**	0.96	124.8
	12-R2	BONCAT**	1.5	150
	12-R3	BONCAT**	1.1	110
	18-R1	BONCAT**	2.26	226

	18-R2	BONCAT[**]	2.02	202
Synthetic feed	18-R3	BONCAT[**]	2.58	258
(positive control) -	24-R1	BONCAT[**]	1.54	154
continued	24-R2	BONCAT[**]	1.5	150
	24-R3	BONCAT[**]	0.895	116.35
	-2	Non-BONCAT[*]	44	4400
	6	Non-BONCAT[*]	60.6	6060
	12	Non-BONCAT[*]	43.8	4380
	18	Non-BONCAT[*]	56.6	5660
	24	Non-BONCAT[*]	66.2	6620
	-2	BONCAT[**]	1.7	170
	6-R1	BONCAT[**]	1.12	112
	6-R2	BONCAT[**]	1.44	144
	6-R3	BONCAT[**]	1.29	129
Water	12-R1	BONCAT[**]	1.11	144.3
(negative control)	12-R2	BONCAT[**]	0.95	123.5
	12-R3	BONCAT[**]	2.08	208
	18-R1	BONCAT[**]	1.18	118
	18-R2	BONCAT[**]	1.89	189
	18-R3	BONCAT[**]	2.54	254
	24-R1	BONCAT[**]	0.835	108.55
	24-R2	BONCAT[**]	2.64	264
	24-R3	BONCAT[**]	6.02	301

Legend: * = Non-BONCAT samples were sent for sequencing using Nanopore (long-reads) and Illumina (short-reads); ** = BONCAT samples were sent for sequencing using Illumina (short-reads); DNA quantification were done using Qubit®.

Appendix D

Table 2. Proteins identified in all treatments after Scaffold filtering (protein 5% FDR; min number of peptides 2; peptide 5% FDR). Proteins were searched against BONCAT, non-BONCAT and contaminants databases.

Substrate	Sample (h)	Grouping	Identified proteins	Identified spectra
Octanoic acid (C$_8$)	-2	Non-BONCAT	44	57
	6	Non-BONCAT	42	62
	12	Non-BONCAT	52	55
	18	Non-BONCAT	58	73
	24	Non-BONCAT	26	14
	-2	BONCAT	52	115
	6	BONCAT	49	65
	12	BONCAT	37	43
	18	BONCAT	57	75
	24	BONCAT	49	60
Synthetic feed (positive control)	12	Non-BONCAT	75	184
	12	BONCAT	86	186
Water (negative control)	12	Non-BONCAT	37	36
	12	BONCAT	53	65
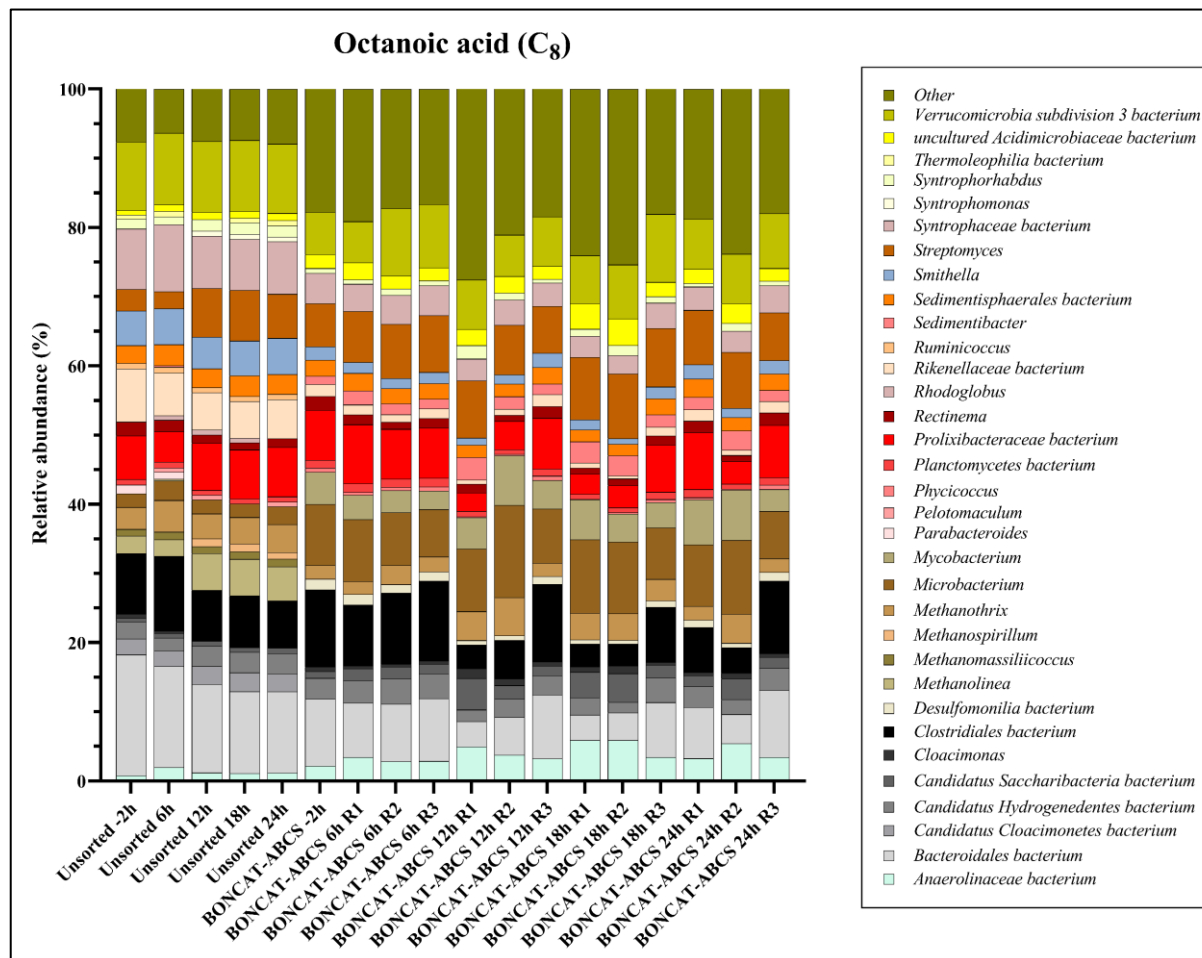
Appendix E. Relative abundance of top-30 most abundant genera based on metagenomic sequencing of sorted and unsorted samples retrieved from octanoic acid fed digesters.
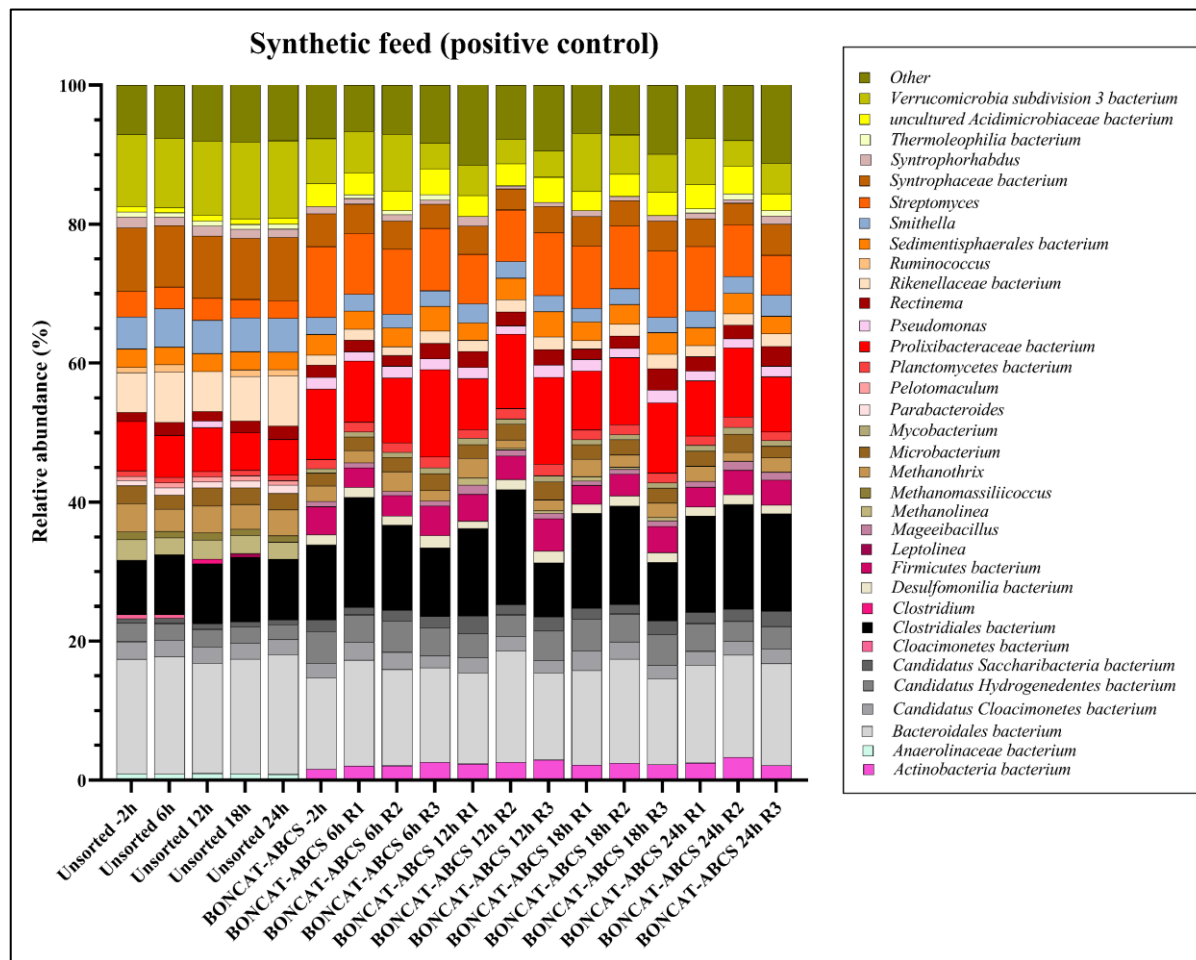
Appendix F. Relative abundance of top-30 most abundant genera based on metagenomic sequencing of sorted and unsorted samples retrieved from nutrient rich synthetic feed digesters.
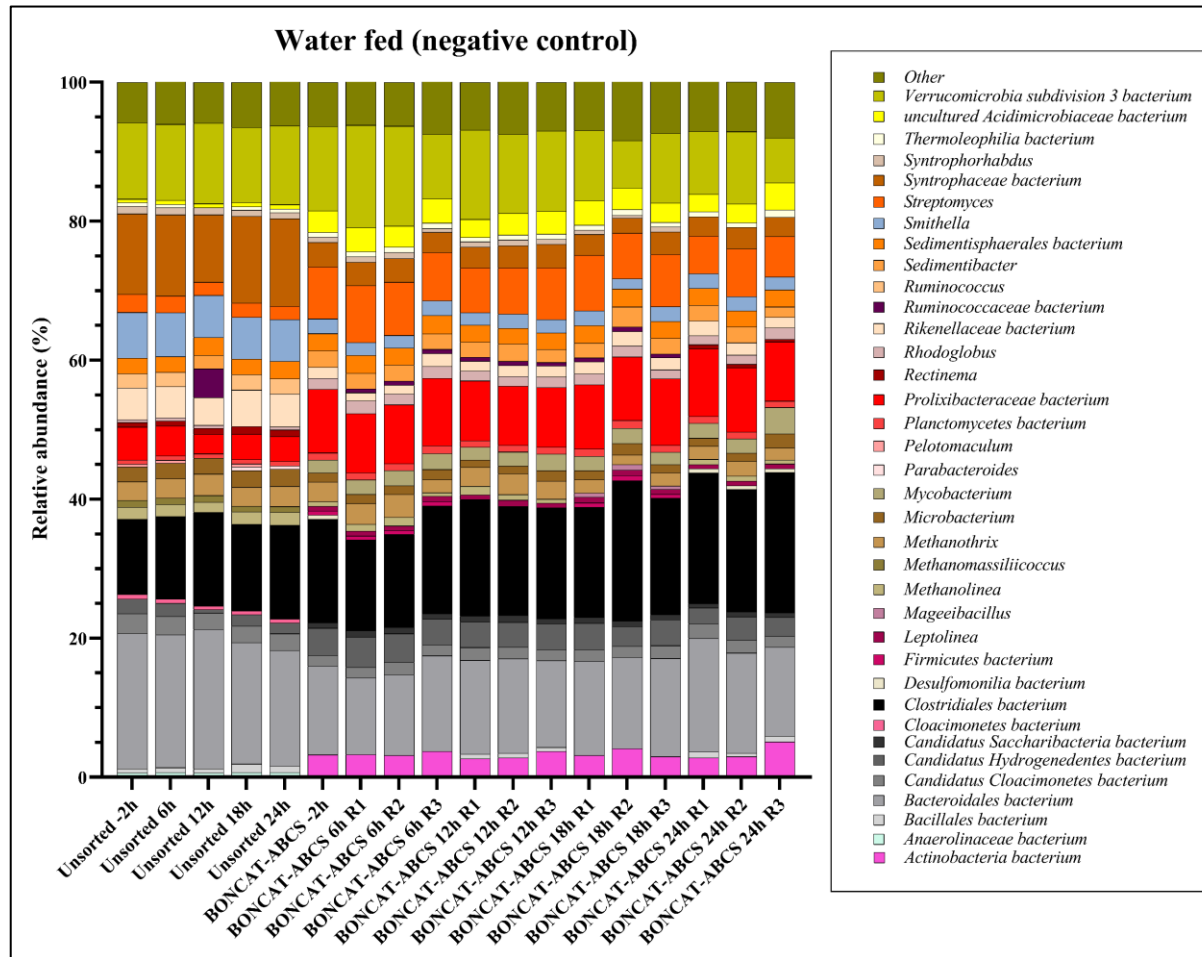
Appendix G. Relative abundance of top-30 most abundant genera based on metagenomic sequencing of sorted and unsorted samples retrieved from water fed digesters.

Appendix H

Table 3 - Complete list of identified proteins per sample prior to Scaffold validation.

	Biological sample	Protein name	Protein accession numbers	Protein molecular weight (Da)	Protein identification probability	Exclusive unique peptide count	Exclusive unique spectrum count	Total spectrum count	Percentage of total spectra	Percentage sequence coverage
1	C8_B_0H	hypothetical protein	ALNLLLBE_641604	28724.7	100%	8	9	14	0.14%	40%
2	C8_B_0H	hypothetical protein	ALNLLLBE_642509	63736.2	100%	8	8	18	0.18%	25%
3	C8_B_0H	hypothetical protein	ALNLLLBE_743383	28982.8	57%	0	0	4	0.04%	9%
4	C8_B_0H	hypothetical protein	ALNLLLBE_197203	28469	78%	1	1	4	0.04%	9%
5	C8_B_0H	hypothetical protein	ALNLLLBE_641606	45516.9	100%	4	4	4	0.04%	12%
6	C8_B_0H	hypothetical protein	ALNLLLBE_641603	61392.1	100%	6	6	6	0.06%	10.00%
7	C8_B_0H	Carbon monoxide dehydrogenase/acetyl-CoA synthase subunit alpha	ALNLLLBE_334547	51772.6	100%	4	4	9	0.09%	11.00%
8	C8_B_0H	hypothetical protein	ALNLLLBE_743321	93204.8	95%	4	4	4	0.04%	5.80%
9	C8_B_0H	hypothetical protein	ALNLLLBE_641420	24057.7	100%	2	2	4	0.04%	15%
10	C8_B_0H	hypothetical protein	ALNLLLBE_334549	89357.7	96%	3	3	6	0.06%	6%
11	C8_B_0H	hypothetical protein	ALNLLLBE_645567	243258.6	100%	2	2	2	0.02%	1.10%
12	C8_B_0H	Elongation factor Tu	ALNLLLBE_246749	43095.6	99%	2	2	4	0.04%	7.60%
13	C8_B_0H	Elongation factor Tu	ALNLLLBE_483213	42931.5	85%	0	0	1	0.01%	4.00%
14	C8_B_0H	Respiratory nitrate reductase 2 beta chain	ALNLLLBE_173240	57135.2	96%	1	1	5	0.05%	4.50%
15	C8_B_0H	Respiratory nitrate reductase 1 beta chain	ALNLLLBE_218625	56752	15%	0	0	2	0.02%	1.40%
16	C8_B_0H	hypothetical protein	ALNLLLBE_641330	76715.7	99%	6	6	7	0.07%	12.00%
17	C8_B_0H	hypothetical protein	ALNLLLBE_643285	53571.2	97%	1	1	4	0.04%	11.00%
18	C8_B_0H	hypothetical protein	ALNLLLBE_438590	57521.6	34%	1	1	1	0.01%	1.10%
19	C8_B_0H	hypothetical protein	ALNLLLBE_580468	78873	75%	2	2	3	0.03%	4.40%
20	C8_B_0H	hypothetical protein	ALNLLLBE_259625	55559	77%	2	2	2	0.02%	6.00%

21	C8_B_0H	hypothetical protein	ALNLLLBE_90799	42809.2	28%	1	1	1	0.01%	0.00%
22	C8_B_0H	hypothetical protein	ALNLLLBE_334511	97269.8	93%	3	3	3	0.03%	4.10%
23	C8_B_0H	hypothetical protein	ALNLLLBE_342545	78943.6	34%	0	0	1	0.01%	2.30%
24	C8_B_0H	hypothetical protein	ALNLLLBE_366056	5654.4	25%	1	1	1	0.01%	18.00%
25	C8_B_0H	hypothetical protein	ALNLLLBE_624953	5654.4	25%	1	1	1	0.01%	18.00%
26	C8_B_0H	hypothetical protein	ALNLLLBE_700955	95628.6	80%	1	1	1	0.01%	1.30%
27	C8_B_0H	hypothetical protein	ALNLLLBE_474867	67684.4	22%	1	1	1	0.01%	1.60%
28	C8_B_0H	hypothetical protein	ALNLLLBE_219451	23245	29%	1	1	1	0.01%	6.00%
29	C8_B_0H	hypothetical protein	ALNLLLBE_642510	53561	10%	0	0	3	0.03%	6.00%
30	C8_B_0H	hypothetical protein	ALNLLLBE_342544	70446	9%	0	0	1	0.01%	2.50%
31	C8_B_0H	hypothetical protein	ALNLLLBE_90712	64996	54%	1	1	1	0.01%	3.00%
32	C8_B_0H	Respiratory nitrate reductase 1 beta chain	ALNLLLBE_584064	57747.9	94%	3	3	3	0.03%	6.90%
33	C8_B_0H	Outer membrane porin protein 32	ALNLLLBE_296280	33703	27%	1	1	1	0.01%	3.00%
34	C8_B_0H	Nitrogen regulatory protein P-II	ALNLLLBE_426853	14100	43%	1	1	1	0.01%	8.60%
35	C8_B_0H	K(+)-insensitive pyrophosphate-energized proton pump	ALNLLLBE_389495	57984	7%	0	0	1	0.01%	1.90%
36	C8_B_0H	K(+)-insensitive pyrophosphate-energized proton pump	ALNLLLBE_11910	70658	7%	0	0	1	0.01%	1.60%
37	C8_B_0H	ATP synthase subunit beta	ALNLLLBE_343885	51090	41%	1	1	2	0.02%	3.00%
38	C8_B_0H	hypothetical protein	ALNLLLBE_342549	41535.4	100%	3	3	3	0.03%	11.00%
39	C8_B_0H	DNA-directed RNA polymerase subunit beta	ALNLLLBE_277117	89607	12%	0	0	1	0.01%	1.70%
40	C8_B_0H	3-hydroxylaminophenol mutase	ALNLLLBE_142484	52267.8	71%	1	1	1	0.01%	3.00%
41	C8_B_0H	60 kDa chaperonin 5	ALNLLLBE_06125	58496	12%	0	0	1	0.01%	1.30%
42	C8_B_0H	60 kDa chaperonin 5	ALNLLLBE_67393	57748.3	100%	2	2	3	0.03%	6.20%
43	C8_B_0H	60 kDa chaperonin 1	ALNLLLBE_399492	37451.1	7%	0	0	1	0.01%	2.00%
44	C8_B_0H	60 kDa chaperonin	ALNLLLBE_333915	57697	8%	0	0	1	0.01%	1.30%
45	C8_B_0H	60 kDa chaperonin	ALNLLLBE_131203	49856	8%	0	0	2	0.02%	5.10%

46	C8_B_0H	60 kDa chaperonin	**ALNLLLBE_583384**	57922.5	5%	0	0	1	0.01%	1.30%
47	C8_B_0H	Cyanate hydratase	**ALNLLLBE_88274**	18009.5	69%	2	2	2	0.02%	11%
48	C8_B_0H	Cold shock protein CspA	**ALNLLLBE_20439**	8611.9	76%	2	2	3	0.03%	20%
49	C8_B_0H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_330937**	46094.9	94%	3	3	5	0.05%	13%
50	C8_B_0H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_369805**	45940.6	94%	3	3	5	0.05%	13%
51	C8_B_0H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_477611**	46315.1	5%	0	0	1	0.01%	4%
52	C8_B_0H	ATP synthase subunit beta 1	**ALNLLLBE_189018**	50810.7	99%	2	2	3	0.03%	7%
53	C8_B_0H	ALNLLLBE_614387-DECOY	**ALNLLLBE_614387-DECOY**		20%	1	1	1	0.01%	
1	C8_B_6H	hypothetical protein	**ALNLLLBE_641604**	28724.7	100%	7	8	12	0.10%	38%
2	C8_B_6H	hypothetical protein	**ALNLLLBE_642509**	63736.2	99%	4	4	5	0.04%	9.50%
3	C8_B_6H	hypothetical protein	**ALNLLLBE_641606**	45516.9	71%	1	1	1	0.08%	2.80%
4	C8_B_6H	hypothetical protein	**ALNLLLBE_641603**	61392.1	100%	5	5	6	0.05%	9%
5	C8_B_6H	hypothetical protein	**ALNLLLBE_641420**	24058	82%	1	1	2	0.02%	8%
6	C8_B_6H	hypothetical protein	**ALNLLLBE_743321**	93204.8	90%	3	3	3	0.03%	4%
7	C8_B_6H	hypothetical protein	**ALNLLLBE_743382**	62548	17%	0	0	1	0.01%	1%
8	C8_B_6H	hypothetical protein	**ALNLLLBE_743383**	28982.8	43%	0	0	3	0.03%	9%
9	C8_B_6H	hypothetical protein	**ALNLLLBE_342549**	41535.4	74%	2	2	3	0.03%	7%
10	C8_B_6H	Carbon monoxide dehydrogenase/acetyl-CoA synthase subunit alpha	**ALNLLLBE_334547**	51772.6	100%	3	3	4	0.03%	9.80%
11	C8_B_6H	hypothetical protein	**ALNLLLBE_334549**	89357.7	95%	2	2	4	0.03%	5.00%
12	C8_B_6H	hypothetical protein	**ALNLLLBE_645567**	243258.6	94%	3	3	3	0.03%	1.70%
13	C8_B_6H	hypothetical protein	**ALNLLLBE_641330**	76715.7	50%	1	1	1	0.01%	1.60%
14	C8_B_6H	hypothetical protein	**ALNLLLBE_641342**	26714.3	82%	2	2	3	0.03%	7.90%
15	C8_B_6H	hypothetical protein	**ALNLLLBE_643285**	53571.2	19%	0	0	2	0.02%	2.60%
16	C8_B_6H	hypothetical protein	**ALNLLLBE_580468**	78873	74%	1	1	2	0.02%	2.80%

17	C8_B_6H	hypothetical protein	**ALNLLLBE_474867**	67684.4	67%	2	2	2	0.02%	2.60%
18	C8_B_6H	hypothetical protein	**ALNLLLBE_197203**	28469	51%	1	1	3	0.03%	9.10%
19	C8_B_6H	hypothetical protein	**ALNLLLBE_45809**	28995.9	39%	1	1	1	0.01%	6.60%
20	C8_B_6H	hypothetical protein	**ALNLLLBE_438590**	57521.6	39%	1	1	1	0.01%	1.10%
21	C8_B_6H	hypothetical protein	**ALNLLLBE_90799**	42809.2	57%	2	2	2	0.02%	0.00%
22	C8_B_6H	hypothetical protein	**ALNLLLBE_105999**	76402.6	44%	1	1	1	0.01%	2.20%
23	C8_B_6H	hypothetical protein	**ALNLLLBE-75826**	83710.8	44%	1	1	1	0.01%	2.00%
24	C8_B_6H	hypothetical protein	**ALNLLLBE_334511**	97269.8	36%	1	1	1	0.01%	1.30%
25	C8_B_6H	hypothetical protein	**ALNLLLBE_54401**	89183	10%	0	0	2	0.02%	1.70%
26	C8_B_6H	hypothetical protein	**ALNLLLBE_761442**	23483	7%	0	0	1	0.01%	5.50%
27	C8_B_6H	hypothetical protein	**ALNLLLBE_642510**	53561	8%	0	0	1	0.01%	1.20%
28	C8_B_6H	hypothetical protein	**ALNLLLBE_90712**	64996	22%	1	1	1	0.01%	3.00%
29	C8_B_6H	Elongation factor Tu	**ALNLLLBE_246749**	43095.6	71%	1	1	1	0.01%	3.50%
30	C8_B_6H	Elongation factor Tu	**ALNLLLBE_203705**	43232.5	28%	1	1	1	0.01%	2.80%
31	C8_B_6H	DNA-directed RNA polymerase subunit beta	**ALNLLLBE_454561**	154237.5	8%	0	0	1	0.01%	0.44%
32	C8_B_6H	Respiratory nitrate reductase 1 beta chain	**ALNLLLBE_584064**	57747.9	86%	2	2	4	0.03%	6.30%
33	C8_B_6H	Respiratory nitrate reductase 1 beta chain	**ALNLLLBE_218625**	56752	16%	0	0	1	0.01%	2.40%
34	C8_B_6H	Respiratory nitrate reductase 2 beta chain	**ALNLLLBE_173240**	57135.2	74%	1	1	3	0.03%	7.30%
35	C8_B_6H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_330937**	46094.9	15%	0	0	1	0.01%	4.00%
36	C8_B_6H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_369805**	45940.6	15%	0	0	1	0.01%	4.00%
37	C8_B_6H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_477611**	46315.1	10%	0	0	1	0.01%	4.00%
38	C8_B_6H	3-hydroxylaminophenol mutase	**ALNLLLBE_142484**	52267.8	80%	2	2	2	0.02%	5.10%
39	C8_B_6H	60 kDa chepronin 5	**ALNLLLBE_67393**	57748.3	93%	2	2	5	0.04%	8.60%
40	C8_B_6H	60 kDa chepronin 5	**ALNLLLBE_06125**	58496	74%	0	0	2	0.02%	3.50%

41	C8_B_6H	60 kDa chepronin 1	ALNLLLBE_399492	37451.1	42%	0	0	2	0.02%	5.80%
42	C8_B_6H	60 kDa chepronin	ALNLLLBE_333915	57697	7%	0	0	1	0.01%	1.30%
43	C8_B_6H	Cyanate hydratase	ALNLLLBE_88274	18009.5	36%	1	1	1	0.01%	5.00%
44	C8_B_6H	Cold shock protein CspA	ALNLLLBE_20439	8612	34%	1	1	1	0.01%	11.00%
1	C8_B_12H	hypothetical protein	ALNLLLBE_641604	28724.7	100%	10	13	16	0.13%	42%
2	C8_B_12H	hypothetical protein	ALNLLLBE_642509	63736.2	100%	11	11	15	0.02%	23%
3	C8_B_12H	hypothetical protein	ALNLLLBE_641603	61392.1	95%	4	5	6	0.05%	10%
4	C8_B_12H	hypothetical protein	ALNLLLBE_743383	28983	97%	4	4	8	0.07%	27%
5	C8_B_12H	hypothetical protein	ALNLLLBE_643285	53571.2	6%	0	0	1	0.01%	1%
6	C8_B_12H	hypothetical protein	ALNLLLBE_641420	24057.7	100%	4	4	6	0.05%	23%
7	C8_B_12H	hypothetical protein	ALNLLLBE_641330	76715.7	92%	5	5	5	0.04%	10.00%
8	C8_B_12H	hypothetical protein	ALNLLLBE_743321	93205	84%	2	2	2	0.02%	2.60%
9	C8_B_12H	hypothetical protein	ALNLLLBE_743386	46186	39%	1	1	1	0.01%	3.20%
10	C8_B_12H	hypothetical protein	ALNLLLBE_675543	67904	69%	2	2	2	0.02%	4.90%
11	C8_B_12H	hypothetical protein	ALNLLLBE_342544	70446	14%	0	0	2	0.02%	7.90%
12	C8_B_12H	hypothetical protein	ALNLLLBE_90799	42809	30%	1	1	1	0.01%	0.00%
13	C8_B_12H	hypothetical protein	ALNLLLBE_422424	40584	83%	2	2	2	0.02%	5.60%
14	C8_B_12H	hypothetical protein	ALNLLLBE_422731	53184	83%	2	2	2	0.02%	4.30%
15	C8_B_12H	Elongation factor Tu	ALNLLLBE_246749	43095.6	93%	2	3	3	0.03%	7%
16	C8_B_12H	Elongation factor Tu	ALNLLLBE_483213	42932	46%	0	0	1	0.01%	4%
17	C8_B_12H	Respiratory nitrate reductase 2 beta chain	ALNLLLBE_173240	57135.2	94%	1	1	5	0.04%	8.90%
18	C8_B_12H	Respiratory nitrate reductase 1 beta chain	ALNLLLBE_218625	56752	51%	0	0	3	0.03%	4.00%
19	C8_B_12H	Carbon monoxide dehydrogenase/acetyl-CoA synthase subunit alpha	ALNLLLBE_334547	51772.6	99%	2	2	4	0.03%	9.80%
20	C8_B_12H	3-hydroxylaminophenol mutase	ALNLLLBE_142484	52267.8	75%	1	1	1	0.01%	2%
21	C8_B_12H	hypothetical protein	ALNLLLBE_342549	41535.4	82%	1	1	1	0.01%	3.10%

22	C8_B_12H	hypothetical protein	**ALNLLLBE_259625**	55559	18%	1	1	1	0.01%	3.30%
23	C8_B_12H	hypothetical protein	**ALNLLLBE_474867**	67684.4	36%	1	1	1	0.01%	1.60%
24	C8_B_12H	hypothetical protein	**ALNLLLBE_641342**	26714	75%	1	1	1	0.01%	5.00%
25	C8_B_12H	hypothetical protein	**ALNLLLBE_761442**	23483	10%	0	0	2	0.02%	6.00%
26	C8_B_12H	hypothetical protein	**ALNLLLBE_105999**	76402.6	53%	1	2	2	0.02%	2.20%
27	C8_B_12H	hypothetical protein	**ALNLLLBE_75826**	83710.8	53%	1	2	2	0.02%	2.00%
28	C8_B_12H	hypothetical protein	**ALNLLLBE_669128**	31780	41%	1	1	2	0.02%	3.50%
29	C8_B_12H	Outer membrane porin protein 32	**ALNLLLBE_296280**	33703	52%	1	1	1	0.01%	3.00%
30	C8_B_12H	60 kDa chaperonin	**ALNLLLBE_121083**	58153	67%	1	1	1	0.08%	3.10%
31	C8_B_12H	60 kDa chaperonin 1	**ALNLLLBE_399492**	37451.1	26%	0	0	1	0.01%	3.80%
32	C8_B_12H	60 kDa chaperonin 5	**ALNLLLBE_67393**	57748.3	34%	0	0	1	0.01%	2.40%
33	C8_B_12H	60 kDa chaperonin 5	**ALNLLLBE_06125**	58496	28%	0	0	1	0.01%	2.20%
34	C8_B_12H	Cyanate hydratase	**ALNLLLBE_88274**	18009.5	66%	2	2	2	0.02%	11.00%
35	C8_B_12H	ATP synthase subunit beta	**ALNLLLBE_376596**	42928	61%	1	1	2	0.02%	6.00%
36	C8_B_12H	ATP synthase subunit beta	**ALNLLLBE_343885**	51090	50%	1	1	1	0.01%	3.00%
37	C8_B_12H	ATP synthase subunit beta	**ALNLLLBE_756712**	54705	70%	1	1	2	0.02%	5.00%
38	C8_B_12H	ATP synthase subunit beta	**ALNLLLBE_441698**	34209	27%	0	0	1	0.01%	4.50%
39	C8_B_12H	hypothetical protein	**ALNLLLBE_641606**	45517	100%	5	5	6	0.05%	15.00%
40	C8_B_12H	hypothetical protein	**ALNLLLBE_197203**	28469	89%	2	2	6	0.05%	13.00%
41	C8_B_12H	hypothetical protein	**ALNLLLBE_334549**	89358	65%	1	1	2	0.02%	2.10%
42	C8_B_12H	hypothetical protein	**ALNLLLBE_645567**	24325	99%	5	5	5	0.04%	2.70%
43	C8_B_12H	hypothetical protein	**ALNLLLBE_54401**	89183	9%	0	0	1	0.01%	0.87%
44	C8_B_12H	hypothetical protein	**ALNLLLBE_584064**	57748	88%	2	2	3	0.03%	7.30%
45	C8_B_12H	hypothetical protein	**ALNLLLBE_743370**	25148	20%	1	1	1	0.01%	5.00%
46	C8_B_12H	hypothetical protein	**ALNLLLBE_743368**	25948	20%	1	1	1	0.01%	4.90%
47	C8_B_12H	hypothetical protein	**ALNLLLBE_466046**	65802	68%	1	1	1	0.01%	1.70%

48	C8_B_12H	hypothetical protein	**ALNLLLBE_466048**	14584	44%	1	1	1	0.01%	9.90%
49	C8_B_12H	Sialic acid-binding periplasmic protein SiaP	**ALNLLLBE_790261**	37481	97%	2	2	2	0.02%	8.10%
50	C8_B_12H	Sialic acid-binding periplasmic protein SiaP	**ALNLLLBE_27688**	37511	97%	2	2	2	0.02%	8.10%
51	C8_B_12H	Sialic acid-binding periplasmic protein SiaP	**ALNLLLBE_310190**	37510	97%	2	2	2	0.02%	8.10%
52	C8_B_12H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_369805**	45941	15%	0	0	1	0.01%	4.00%
53	C8_B_12H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_330937**	46095	15%	0	0	1	0.01%	4.00%
54	C8_B_12H	V-type ATP synthase beta chain	**ALNLLLBE_644750**	50785	27%	0	0	1	0.01%	3.00%
55	C8_B_12H	V-type ATP synthase alpha chain	**ALNLLLBE_54428**	63580	22%	1	1	1	0.01%	2.30%
56	C8_B_12H	V-type ATP synthase beta chain	**ALNLLLBE_54427**	50993	6%	0	0	1	0.01%	3.00%
57	C8_B_12H	Formate dehydrogenase H	**ALNLLLBE_743668**	76707	43%	1	1	1	0.01%	1.60%
58	C8_B_12H	hypothetical protein	**ALNLLLBE_342545**	78944	69%	1	1	3	0.03%	12.00%
59	C8_B_12H	hypothetical protein	**ALNLLLBE_366056**	5654	48%	1	1	1	0.01%	24.00%
60	C8_B_12H	hypothetical protein	**ALNLLLBE_624953**	5654	48%	1	1	1	0.01%	24.00%
61	C8_B_12H	Glycerol kinase	**ALNLLLBE_232721**	55015	14%	0	0	1	0.01%	2.00%
62	C8_B_12H	Glycerol kinase	**ALNLLLBE_395413**	55022	60%	1	1	2	0.02%	4.00%
63	C8_B_12H	hypothetical protein	**ALNLLLBE_642510**	53561	32%	1	1	1	0.01%	2.30%
64	C8_B_12H	K(+)-insensitive pyrophosphate-energized proton pump	**ALNLLLBE_389495**	57984	23%	0	0	1	0.01%	1.90%
65	C8_B_12H	K(+)-insensitive pyrophosphate-energized proton pump	**ALNLLLBE_11910**	70658	23%	0	0	1	0.01%	1.60%
66	C8_B_12H	Acetyl-coenzyme A synthetase	**ALNLLLBE_642943**	75311	70%	1	1	1	0.01%	1.80%
67	C8_B_12H	NAD-specific glutamate dehydrogenase	**ALNLLLBE_791579**	48866	16%	1	1	1	0.01%	4.00%
68	C8_B_12H	hypothetical protein	**ALNLLLBE_90712**	64996	30%	1	1	1	0.01%	3.00%
69	C8_B_12H	Methylcorrinoid:tetrahydrofolate methyltransferase	**ALNLLLBE_645360**	33749	82%	2	2	2	0.02%	8.00%
70	C8_B_12H	Methylcorrinoid:tetrahydrofolate methyltransferase	**ALNLLLBE_92794**	33812	82%	2	2	2	0.02%	8.00%

71	C8_B_12H	50S ribosomal protein L1	ALNLLLBE_246746	23966	33%	1	1	1	0.01%	5.60%
72	C8_B_12H	hypothetical protein	ALNLLLBE_756266	10428	50%	2	2	2	0.02%	1.20%
1	C8_B_18H	hypothetical protein	ALNLLLBE_641604	28724.7	100%	6	6	8	0.11%	31%
2	C8_B_18H	hypothetical protein	ALNLLLBE_642509	63736.2	100%	3	3	3	0.04%	6.10%
3	C8_B_18H	hypothetical protein	ALNLLLBE_641606	45516.9	91%	2	2	2	0.03%	7%
4	C8_B_18H	hypothetical protein	ALNLLLBE_641603	61392.1	98%	2	2	3	0.04%	5%
5	C8_B_18H	hypothetical protein	ALNLLLBE_45809	28995.9	98%	2	2	3	0.04%	9%
6	C8_B_18H	hypothetical protein	ALNLLLBE_342549	42535.4	71%	1	1	1	0.01%	3%
7	C8_B_18H	hypothetical protein	ALNLLLBE_438590	57521.6	38%	1	1	1	0.01%	1%
8	C8_B_18H	hypothetical protein	ALNLLLBE_474867	67684.4	58%	1	1	1	0.01%	2%
9	C8_B_18H	Carbon monoxide dehydrogenase/acetyl-CoA synthase subunit alpha	ALNLLLBE_334557	51772.6	100%	4	4	7	0.10%	18%
10	C8_B_18H	hypothetical protein	ALNLLLBE_743321	93204.8	3500%	1	1	1	0.01%	1.60%
11	C8_B_18H	hypothetical protein	ALNLLLBE_641420	24057.7	75%	1	1	1	0.01%	4.90%
12	C8_B_18H	hypothetical protein	ALNLLLBE_743382	62548	98%	1	1	2	0.03%	3.20%
13	C8_B_18H	hypothetical protein	ALNLLLBE_334549	89357.7	99%	2	2	4	0.05%	5.00%
14	C8_B_18H	hypothetical protein	ALNLLLBE_743383	28982.8	91%	1	1	9	0.12%	14.00%
15	C8_B_18H	hypothetical protein	ALNLLLBE_197203	28469	99%	3	3	10	0.14%	16.00%
16	C8_B_18H	hypothetical protein	ALNLLLBE_54401	89183	32%	0	0	2	0.03%	1.70%
17	C8_B_18H	hypothetical protein	ALNLLLBE_259625	55559	82%	2	2	2	0.03%	6.00%
18	C8_B_18H	hypothetical protein	ALNLLLBE_743368	25947.6	95%	2	2	3	0.04%	11.00%
19	C8_B_18H	hypothetical protein	ALNLLLBE_743370	25148.5	95%	2	2	3	0.04%	11.00%
20	C8_B_18H	hypothetical protein	ALNLLLBE_743386	46186.3	97%	2	2	2	0.03%	3.70%
21	C8_B_18H	hypothetical protein	ALNLLLBE_761442	23483	9%	0	0	2	0.03%	5.50%
22	C8_B_18H	hypothetical protein	ALNLLLBE_366056	5654.4	97%	2	2	2	0.03%	41.00%
23	C8_B_18H	hypothetical protein	ALNLLLBE_624953	5654.4	97%	2	2	2	0.03%	41.00%

24	C8_B_18H	hypothetical protein	**ALNLLLBE_669128**	31779.5	37%	1	1	2	0.03%	3.50%
25	C8_B_18H	hypothetical protein	**ALNLLLBE_466048**	14584	52%	1	1	1	0.01%	9.90%
26	C8_B_18H	hypothetical protein	**ALNLLLBE_753512**	12463	16%	1	1	1	0.01%	6.70%
27	C8_B_18H	Elongation factor Tu	**ALNLLLBE_246749**	43095.6	92%	2	2	3	0.04%	8.80%
28	C8_B_18H	Elongation factor Tu	**ALNLLLBE_483213**	42931.5	71%	0	0	2	0.03%	4.00%
29	C8_B_18H	Elongation factor Tu	**ALNLLLBE_716068**	42735	99%	2	2	2	0.03%	6.40%
30	C8_B_18H	Elongation factor Tu	**ALNLLLBE_203705**	43232.5	99%	1	1	3	0.04%	7.30%
31	C8_B_18H	Elongation factor Tu	**ALNLLLBE_592500**	42979	35%	0	0	2	0.03%	4.50%
32	C8_B_18H	Respiratory nitrate reductase 2 beta chain	**ALNLLLBE_173240**	57135.2	97%	1	1	5	0.07%	8.70%
33	C8_B_18H	Respiratory nitrate reductase 1 beta chain	**ALNLLLBE_218625**	56752	39%	0	0	3	0.04%	3.80%
34	C8_B_18H	Respiratory nitrate reductase 1 beta chain	**ALNLLLBE_584064**	57747.9	76%	2	2	3	0.04%	7.30%
35	C8_B_18H	Leu/Ille/Val-binding protein	**ALNLLLBE_328832**	41182.7	65%	2	2	2	0.03%	5.70%
36	C8_B_18H	Leu/Ille/Val-binding protein	**ALNLLLBE_343029**	41248.7	65%	2	2	2	0.03%	5.70%
37	C8_B_18H	Glutamine synthetase	**ALNLLLBE_315415**	30249	86%	2	2	2	0.03%	8.20%
38	C8_B_18H	hypothetical protein	**ALNLLLBE_641330**	76715.7	31%	1	1	1	0.01%	1.60%
39	C8_B_18H	Phthiodiolone/phenolphthiodiolone dimycocerosates ketoreductase	**ALNLLLBE_743328**	34701	27%	1	1	1	0.01%	3.00%
40	C8_B_18H	Outer membrane porin protein 32	**ALNLLLBE_296280**	33703	44%	1	1	1	0.01%	3.00%
41	C8_B_18H	NAD-specific glutamate dehydrogenase	**ALNLLLBE_791579**	48866	82%	2	2	2	0.03%	6.50%
42	C8_B_18H	K(+)-insensitive pyrophosphate-energized proton pump	**ALNLLLBE_389495**	57984	21%	0	0	1	0.01%	1.90%
43	C8_B_18H	K(+)-insensitive pyrophosphate-energized proton pump	**ALNLLLBE_11910**	70658	21%	0	0	1	0.01%	1.60%
44	C8_B_18H	Formate dehydrogenase H	**ALNLLLBE_743668**	76707	17%	1	1	1	0.01%	1.30%
45	C8_B_18H	3-hydroxylaminophenol mutase	**ALNLLLBE_142484**	52267.8	83%	1	1	1	0.01%	3.00%
46	C8_B_18H	ATP synthase subunit beta	**ALNLLLBE_376596**	42928	69%	1	1	2	0.03%	6.00%
47	C8_B_18H	ATP synthase subunit beta	**ALNLLLBE_343885**	51090	24%	1	1	1	0.01%	3.00%

48	C8_B_18H	ATP synthase subunit beta	**ALNLLLBE_756712**	54705	9%	0	0	1	0.01%	2.80%
49	C8_B_18H	ATP synthase subunit beta	**ALNLLLBE_441698**	34208.7	9%	0	0	1	0.01%	4.50%
50	C8_B_18H	DNA-directed RNA-polymerase subunit beta	**ALNLLLBE_454561**	154237.5	54%	1	1	2	0.03%	1.20%
51	C8_B_18H	Cold shock protein CspA	**ALNLLLBE_20439**	8611.9	70%	1	1	1	0.01%	11.00%
52	C8_B_18H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_330937**	46094.9	89%	1	1	3	0.04%	7%
53	C8_B_18H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_369805**	45940.6	89%	1	1	3	0.04%	7%
54	C8_B_18H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_477611**	46315.1	65%	1	1	3	0.04%	6%
55	C8_B_18H	ALNLLLBE_323811-DECOY	**ALNLLLBE_323811-DECOY**		17%	1	1	1	0.01%	
56	C8_B_18H	60 kDa chaperonin 5	**ALNLLLBE_67393**	57748.3	100%	2	2	5	0.07%	9.70%
57	C8_B_18H	60 kDa chaperonin 5	**ALNLLLBE_06125**	58496	95%	0	0	2	0.03%	3.50%
58	C8_B_18H	60 kDa chaperonin 1	**ALNLLLBE_399492**	37451.1	60%	0	0	3	0.04%	7.50%
59	C8_B_18H	60 kDa chaperonin	**ALNLLLBE_333915**	57697	9%	0	0	1	0.01%	1.30%
60	C8_B_18H	60 kDa chaperonin	**ALNLLLBE_121083**	58153	5%	0	0	1	0.01%	1.30%
1	C8_B_24H	hypothetical protein	**ALNLLLBE_641604**	28724.7	100%	4	4	5	0.06%	26%
2	C8_B_24H	hypothetical protein	**ALNLLLBE_642509**	63736.2	91%	3	3	4	0.05%	7.60%
3	C8_B_24H	hypothetical protein	**ALNLLLBE_641606**	45516.9	51%	1	1	1	0.01%	3%
4	C8_B_24H	hypothetical protein	**ALNLLLBE_641603**	61392.1	98%	3	3	4	0.05%	6%
5	C8_B_24H	hypothetical protein	**ALNLLLBE_45809**	28995.9	86%	2	2	2	0.03%	9%
6	C8_B_24H	hypothetical protein	**ALNLLLBE_438590**	57521.6	34%	1	1	1	0.01%	1%
7	C8_B_24H	hypothetical protein	**ALNLLLBE_474867**	67684.4	59%	1	1	1	0.01%	2%
8	C8_B_24H	Carbon monoxide dehydrogenase/acetyl-CoA synthase subunit alpha	**ALNLLLBE_334557**	51772.6	100%	4	4	5	0.06%	12%
9	C8_B_24H	hypothetical protein	**ALNLLLBE_743321**	93204.8	56%	1	1	1	0.01%	1.60%
10	C8_B_24H	hypothetical protein	**ALNLLLBE_641420**	24057.7	34%	1	1	1	0.01%	4.90%
11	C8_B_24H	hypothetical protein	**ALNLLLBE_743382**	62548	98%	1	1	3	0.04%	3.20%

12	C8_B_24H	hypothetical protein	**ALNLLLBE_334549**	89357.7	97%	1	1	4	0.05%	4.50%
13	C8_B_24H	hypothetical protein	**ALNLLLBE_743383**	28982.8	83%	1	1	8	0.10%	14.00%
14	C8_B_24H	hypothetical protein	**ALNLLLBE_197203**	28469	87%	2	2	8	0.10%	13.00%
15	C8_B_24H	hypothetical protein	**ALNLLLBE_259625**	55559	93%	2	2	2	0.03%	6.00%
16	C8_B_24H	hypothetical protein	**ALNLLLBE_743368**	25947.6	98%	2	2	2	0.03%	11.00%
17	C8_B_24H	hypothetical protein	**ALNLLLBE_743370**	25148.5	98%	2	2	2	0.03%	11.00%
18	C8_B_24H	hypothetical protein	**ALNLLLBE_743386**	46186.3	64%	1	1	1	0.01%	1.60%
19	C8_B_24H	hypothetical protein	**ALNLLLBE_669128**	31779.5	62%	2	2	4	0.05%	3.90%
20	C8_B_24H	hypothetical protein	**ALNLLLBE_54401**	89183	41%	0	0	3	0.04%	2.90%
21	C8_B_24H	hypothetical protein	**ALNLLLBE_219451**	23245	23%	1	1	1	0.01%	6.00%
22	C8_B_24H	Elongation factor Tu	**ALNLLLBE_246749**	43095.6	100%	3	3	3	0.04%	11.00%
23	C8_B_24H	Elongation factor Tu	**ALNLLLBE_716068**	42735	60%	1	1	1	0.01%	2.30%
24	C8_B_24H	Elongation factor Tu	**ALNLLLBE_203705**	43232.5	85%	1	1	1	0.01%	2.80%
25	C8_B_24H	Respiratory nitrate reductase 2 beta chain	**ALNLLLBE_173240**	57135.2	93%	1	1	1	0.01%	3.20%
26	C8_B_24H	Respiratory nitrate reductase 1 beta chain	**ALNLLLBE_584064**	57747.9	16%	1	1	1	0.01%	3.10%
27	C8_B_24H	Outer membrane porin protein 32	**ALNLLLBE_296280**	33703	29%	1	1	1	0.01%	3.00%
28	C8_B_24H	Nitrogen regulatory protein P-II	**ALNLLLBE_426853**	14100	60%	2	2	2	0.03%	22.00%
29	C8_B_24H	Leu/Ille/Val-binding protein	**ALNLLLBE_328832**	41182.7	39%	1	1	1	0.01%	1.80%
30	C8_B_24H	Leu/Ille/Val-binding protein	**ALNLLLBE_343029**	41248.7	39%	1	1	1	0.01%	1.80%
31	C8_B_24H	K(+)-insensitive pyrophosphate-energized proton pump	**ALNLLLBE_11910**	70658	7%	0	0	1	0.01%	1.60%
32	C8_B_24H	K(+)-insensitive pyrophosphate-energized proton pump	**ALNLLLBE_389495**	57984	7%	0	0	1	0.01%	1.90%
33	C8_B_24H	3-hydroxylaminophenol mutase	**ALNLLLBE_142484**	52267.8	96%	1	1	1	0.01%	3.00%
34	C8_B_24H	DNA-directed RNA-polymerase subunit beta	**ALNLLLBE_277117**	89607	33%	0	0	1	0.01%	1.10%
35	C8_B_24H	DNA-directed RNA-polymerase subunit beta	**ALNLLLBE_550114**	38019	16%	0	0	1	0.01%	2.60%

36	C8_B_24H	DNA-directed RNA-polymerase subunit beta	**ALNLLLBE_51534**	143539	62%	1	1	2	0.03%	1.30%
37	C8_B_24H	DNA-directed RNA-polymerase subunit beta	**ALNLLLBE_563289**	143589	62%	1	1	2	0.03%	1.30%
38	C8_B_24H	DNA-directed RNA-polymerase subunit beta	**ALNLLLBE_454561**	154238	5%	0	0	1	0.01%	0.44%
39	C8_B_24H	Cold shock protein CspA	**ALNLLLBE_20439**	8612	29%	1	1	1	0.01%	8.90%
40	C8_B_24H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_330937**	46094.9	94%	3	3	3	0.04%	9%
41	C8_B_24H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_369805**	45940.6	94%	3	3	3	0.04%	9%
42	C8_B_24H	60 kDa chaperonin 5	**ALNLLLBE_67393**	57748.3	99%	1	1	3	0.04%	5.50%
43	C8_B_24H	60 kDa chaperonin 5	**ALNLLLBE_06125**	58496	93%	0	0	2	0.03%	3.50%
44	C8_B_24H	60 kDa chaperonin 1	**ALNLLLBE_399492**	37451.1	90%	1	1	3	0.04%	9.20%
45	C8_B_24H	60 kDa chaperonin	**ALNLLLBE_333915**	57697	69%	1	1	2	0.03%	3.10%
46	C8_B_24H	60 kDa chaperonin	**ALNLLLBE_583384**	57865	14%	1	1	2	0.03%	3.10%
47	C8_B_24H	60 kDa chaperonin	**ALNLLLBE_121083**	58153	98%	1	1	2	0.03%	4.40%
48	C8_B_24H	ALNLLLBE_614387-DECOY	**ALNLLLBE_614387-DECOY**		15%	1	1	1		
1	CNEG_B_12H	hypothetical protein	**ALNLLLBE_641604**	28724.7	100%	5	5	6	0.07%	30%
2	CNEG_B_12H	hypothetical protein	**ALNLLLBE_642509**	63736.2	92%	3	3	4	0.05%	6.80%
3	CNEG_B_12H	hypothetical protein	**ALNLLLBE_641606**	45516.9	79%	1	1	1	0.01%	3%
4	CNEG_B_12H	hypothetical protein	**ALNLLLBE_641603**	61392.1	98%	2	2	2	0.02%	5%
5	CNEG_B_12H	hypothetical protein	**ALNLLLBE_45809**	28995.9	95%	1	1	2	0.02%	7%
6	CNEG_B_12H	hypothetical protein	**ALNLLLBE_438590**	57521.6	40%	1	1	1	0.01%	1%
7	CNEG_B_12H	hypothetical protein	**ALNLLLBE_474867**	67684.4	91%	2	2	2	0.02%	3%
8	CNEG_B_12H	Carbon monoxide dehydrogenase/acetyl-CoA synthase subunit alpha	**ALNLLLBE_334557**	51772.6	100%	3	3	5	0.06%	11%
9	CNEG_B_12H	hypothetical protein	**ALNLLLBE_641420**	24057.7	84%	1	1	1	0.01%	4.90%
10	CNEG_B_12H	hypothetical protein	**ALNLLLBE_743382**	62548	99%	1	1	1	0.01%	1.80%
11	CNEG_B_12H	hypothetical protein	**ALNLLLBE_334549**	89357.7	99%	3	3	6	0.07%	6.50%

12	CNEG_B_12H	hypothetical protein	**ALNLLLBE_743383**	28982.8	95%	1	1	9	0.10%	14.00%
13	CNEG_B_12H	hypothetical protein	**ALNLLLBE_197203**	28469	96%	2	2	9	0.10%	13.00%
14	CNEG_B_12H	hypothetical protein	**ALNLLLBE_643285**	53571.2	6%	0	0	1	0.01%	1.40%
15	CNEG_B_12H	hypothetical protein	**ALNLLLBE_645567**	243258.6	42%	1	1	1	0.01%	0.72%
16	CNEG_B_12H	hypothetical protein	**ALNLLLBE_743368**	25947.6	100%	3	3	3	0.03%	14.00%
17	CNEG_B_12H	hypothetical protein	**ALNLLLBE_743370**	25148	100%	3	3	3	0.03%	14.00%
18	CNEG_B_12H	hypothetical protein	**ALNLLLBE_743386**	46186.3	78%	2	2	2	0.02%	3.70%
19	CNEG_B_12H	hypothetical protein	**ALNLLLBE_90799**	42809.2	55%	2	2	2	0.02%	0.00%
20	CNEG_B_12H	hypothetical protein	**ALNLLLBE_521398**	142013.7	42%	1	1	1	0.01%	0.95%
21	CNEG_B_12H	hypothetical protein	**ALNLLLBE_669128**	31779.5	75%	2	2	3	0.03%	3.90%
22	CNEG_B_12H	hypothetical protein	**ALNLLLBE_753512**	12463	23%	1	1	1	0.01%	6.70%
23	CNEG_B_12H	hypothetical protein	**ALNLLLBE_761442**	23483	9%	0	0	1	0.01%	5.50%
24	CNEG_B_12H	hypothetical protein	**ALNLLLBE_54401**	89183	31%	0	0	2	0.02%	2.00%
25	CNEG_B_12H	Elongation factor Tu	**ALNLLLBE_246749**	43095.6	96%	1	1	1	0.01%	3.50%
26	CNEG_B_12H	Elongation factor Tu	**ALNLLLBE_483213**	42931.5	85%	0	0	1	0.01%	4.00%
27	CNEG_B_12H	Elongation factor Tu	**ALNLLLBE_526058**	43715.6	68%	2	2	2	0.02%	5.70%
28	CNEG_B_12H	Respiratory nitrate reductase 2 beta chain	**ALNLLLBE_173240**	57135.2	97%	1	1	2	0.02%	5.50%
29	CNEG_B_12H	Respiratory nitrate reductase 1 beta chain	**ALNLLLBE_218625**	56752	21%	0	0	1	0.01%	2.40%
30	CNEG_B_12H	Respiratory nitrate reductase 1 beta chain	**ALNLLLBE_584064**	57747.9	20%	1	1	1	0.01%	3.10%
31	CNEG_B_12H	Phthiodiolone/phenolphthiodiolone dimycocerosates ketoreductase	**ALNLLLBE_743328**	34701	22%	1	1	1	0.01%	3.00%
32	CNEG_B_12H	NAD-specific glutamate dehydrogenase	**ALNLLLBE_791579**	48866.4	73%	2	2	2	0.02%	6.50%
33	CNEG_B_12H	Leu/Ille/Val-binding protein	**ALNLLLBE_328832**	41182.7	30%	1	1	1	0.01%	1.80%
34	CNEG_B_12H	Leu/Ille/Val-binding protein	**ALNLLLBE_343029**	41248.7	30%	1	1	1	0.01%	1.80%
35	CNEG_B_12H	Glutamine synthetase	**ALNLLLBE_315415**	30249	22%	0	0	1	0.01%	5.90%
36	CNEG_B_12H	Formate dehydrogenase H	**ALNLLLBE_743668**	76707	92%	2	2	2	0.02%	2.90%

#	Sample	Protein	Accession							
37	CNEG_B_12H	3-hydroxylaminophenol mutase	**ALNLLLBE_142484**	52267.8	66%	1	1	1	0.01%	3.00%
38	CNEG_B_12H	ATP synthase subunit beta 1	**ALNLLLBE_189018**	50810.7	25%	1	1	1	0.01%	4.10%
39	CNEG_B_12H	ATP synthase subunit beta	**ALNLLLBE_376596**	42928	35%	1	1	1	0.01%	2.50%
40	CNEG_B_12H	ATP synthase subunit alpha	**ALNLLLBE_413871**	53611	59%	1	1	1	0.01%	1.80%
41	CNEG_B_12H	DNA-directed RNA-polymerase subunit beta	**ALNLLLBE_277117**	89607	47%	0	0	1	0.01%	1.10%
42	CNEG_B_12H	DNA-directed RNA-polymerase subunit beta	**ALNLLLBE_372517**	119264	5%	0	0	1	0.01%	0.84%
43	CNEG_B_12H	DNA-directed RNA-polymerase subunit beta	**ALNLLLBE_550114**	38019	24%	0	0	1	0.01%	2.60%
44	CNEG_B_12H	DNA-directed RNA-polymerase subunit beta	**ALNLLLBE_454561**	154238	7%	0	0	1	0.01%	0.44%
45	CNEG_B_12H	DNA-directed RNA-polymerase subunit beta	**ALNLLLBE_310308**	71774	44%	1	1	2	0.02%	2.50%
46	CNEG_B_12H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_330937**	46094.9	82%	1	1	2	0.02%	7%
47	CNEG_B_12H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_369805**	45940.6	82%	1	1	2	0.02%	7%
48	CNEG_B_12H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_477611**	46315.1	6%	0	0	1	0.01%	4%
49	CNEG_B_12H	ALNLLLBE_614387-DECOY	**ALNLLLBE_614387-DECOY**		29%	1	1	1	0.01%	
50	CNEG_B_12H	ALNLLLBE_323811-DECOY	**ALNLLLBE_323811-DECOY**		60%	2	2	2	0.02%	
51	CNEG_B_12H	60 kDa chaperonin 5	**ALNLLLBE_67393**	57748.3	100%	2	2	3	0.03%	6.20%
52	CNEG_B_12H	60 kDa chaperonin 5	**ALNLLLBE_06125**	58496	96%	0	0	2	0.02%	3.50%
53	CNEG_B_12H	60 kDa chaperonin 1	**ALNLLLBE_399492**	37451.1	10%	0	0	1	0.01%	2.00%
54	CNEG_B_12H	60 kDa chaperonin	**ALNLLLBE_333915**	57697	11%	0	0	1	0.01%	1.30%
55	CNEG_B_12H	60 kDa chaperonin	**ALNLLLBE_121083**	58153	7%	0	0	1	0.01%	1.30%
1	CPLUS_B_12H	sn-glycerol-3-phosphate-binding periplasmic protein UgpB	**ALNLLLBE_380224**	48280	96%	4	5	5	0.04%	8%
2	CPLUS_B_12H	sn-glycerol-3-phosphate-binding periplasmic protein UgpB	**ALNLLLBE_380230**	28583	96%	4	5	5	0.04%	13%
3	CPLUS_B_12H	sn-glycerol-3-phosphate-binding periplasmic protein UgpB	**ALNLLLBE_380230**	28595	96%	4	5	5	0.04%	13%
4	CPLUS_B_12H	putative protein	**ALNLLLBE_110201**	35565	99%	5	5	5	0.04%	15%

5	CPLUS_B_12H	putative protein	ALNLLLBE_162024	35551	99%	5	5	5	0.04%	15%
6	CPLUS_B_12H	hypothetical protein	ALNLLLBE_641604	28724.7	100%	11	13	16	0.13%	45%
7	CPLUS_B_12H	hypothetical protein	ALNLLLBE_642509	63736.2	100%	11	11	17	0.14%	28%
8	CPLUS_B_12H	hypothetical protein	ALNLLLBE_641606	45516.9	100%	5	5	5	0.04%	13%
9	CPLUS_B_12H	hypothetical protein	ALNLLLBE_641603	61392.1	100%	5	6	6	0.05%	10%
10	CPLUS_B_12H	hypothetical protein	ALNLLLBE_643285	53571.2	62%	1	1	2	0.02%	4%
11	CPLUS_B_12H	hypothetical protein	ALNLLLBE_743321	93204.8	100%	9	9	9	0.07%	12%
12	CPLUS_B_12H	hypothetical protein	ALNLLLBE_641420	24057.7	99%	3	4	4	0.03%	14%
13	CPLUS_B_12H	hypothetical protein	ALNLLLBE_743382	62548	95%	2	2	2	0.02%	3.00%
14	CPLUS_B_12H	hypothetical protein	ALNLLLBE_334549	89357.7	69%	1	1	2	0.02%	2.50%
15	CPLUS_B_12H	hypothetical protein	ALNLLLBE_743383	28982.8	99%	5	5	6	0.05%	30.00%
16	CPLUS_B_12H	hypothetical protein	ALNLLLBE_645567	243258.6	96%	4	4	4	0.03%	2.20%
17	CPLUS_B_12H	hypothetical protein	ALNLLLBE_641330	76715.7	64%	1	1	1	0.01%	1.60%
18	CPLUS_B_12H	hypothetical protein	ALNLLLBE_197203	28469	41%	1	1	2	0.02%	9.10%
19	CPLUS_B_12H	hypothetical protein	ALNLLLBE_54401	89183	9%	0	0	1	0.01%	0.87%
20	CPLUS_B_12H	hypothetical protein	ALNLLLBE_474867	67684	39%	1	1	1	0.01%	1.60%
21	CPLUS_B_12H	hypothetical protein	ALNLLLBE_743386	46186	67%	3	3	3	0.02%	6.40%
22	CPLUS_B_12H	hypothetical protein	ALNLLLBE_466046	65802	65%	1	1	1	0.01%	1.70%
23	CPLUS_B_12H	hypothetical protein	ALNLLLBE_641342	26714	79%	2	2	2	0.02%	7.90%
24	CPLUS_B_12H	hypothetical protein	ALNLLLBE_761442	23483	13%	0	0	2	0.02%	6.00%
25	CPLUS_B_12H	hypothetical protein	ALNLLLBE_669128	31780	30%	1	1	1	0.01%	3.50%
26	CPLUS_B_12H	hypothetical protein	ALNLLLBE_624953	5654	84%	1	1	1	0.01%	24.00%
27	CPLUS_B_12H	hypothetical protein	ALNLLLBE_366056	5654	84%	1	1	1	0.01%	24.00%
28	CPLUS_B_12H	hypothetical protein	ALNLLLBE_642510	53561	58%	1	1	3	0.02%	7.00%
29	CPLUS_B_12H	hypothetical protein	ALNLLLBE_653168	99054	17%	0	0	1	0.01%	1.40%
30	CPLUS_B_12H	hypothetical protein	ALNLLLBE_90712	64996	34%	1	1	1	0.01%	3.00%

#	Sample	Description	Accession							
31	CPLUS_B_12H	hypothetical protein	ALNLLLBE_744241	69352	72%	2	2	2	0.02%	3.70%
32	CPLUS_B_12H	hypothetical protein	ALNLLLBE_764319	96026	73%	0	0	3	0.02%	3.70%
33	CPLUS_B_12H	hypothetical protein	ALNLLLBE_763559	101797	59%	0	0	3	0.02%	3.00%
34	CPLUS_B_12H	hypothetical protein	ALNLLLBE_744739	29925	70%	2	2	2	0.02%	8.00%
35	CPLUS_B_12H	hypothetical protein	ALNLLLBE_21874	92364	79%	3	3	3	0.02%	5.30%
36	CPLUS_B_12H	hypothetical protein	ALNLLLBE_756266	104280	21%	1	1	1	0.01%	0.62%
37	CPLUS_B_12H	hypothetical protein	ALNLLLBE_80641	47532	64%	2	2	2	0.02%	6.00%
38	CPLUS_B_12H	Elongation factor Tu	ALNLLLBE_246749	43095.6	80%	1	2	2	0.02%	3%
39	CPLUS_B_12H	Elongation factor Tu	ALNLLLBE_203705	43233	34%	0	0	1	0.01%	2%
40	CPLUS_B_12H	Elongation factor Tu	ALNLLLBE_592500	42978.8	22%	0	0	1	0.01%	2%
41	CPLUS_B_12H	Carbon monoxide dehydrogenase/acetyl-CoA synthase subunit alpha	ALNLLLBE_334547	51772.6	100%	3	3	5	0.04%	12.00%
42	CPLUS_B_12H	Respiratory nitrate reductase 1 beta chain	ALNLLLBE_584064	57747.9	97%	3	3	4	0.03%	8.60%
43	CPLUS_B_12H	Respiratory nitrate reductase 2 beta chain	ALNLLLBE_173240	57135	98%	1	1	5	0.04%	8.90%
44	CPLUS_B_12H	hypothetical protein	ALNLLLBE_521398	142013.7	79%	2	2	2	0.02%	2.10%
45	CPLUS_B_12H	hypothetical protein	ALNLLLBE_342545	78943.6	89%	0	0	3	0.02%	8.90%
46	CPLUS_B_12H	hypothetical protein	ALNLLLBE_342549	41535.4	99%	4	4	4	0.03%	16.00%
47	CPLUS_B_12H	hypothetical protein	ALNLLLBE_733678	39382	97%	5	5	6	0.05%	17.00%
48	CPLUS_B_12H	Glycerol kinase	ALNLLLBE_232721	55014.5	100%	4	4	7	0.06%	13.00%
49	CPLUS_B_12H	Glycerol kinase	ALNLLLBE_395413	55022	98%	3	3	6	0.05%	11.00%
50	CPLUS_B_12H	Glycerol kinase	ALNLLLBE_772310	29102	35%	1	1	2	0.02%	6.90%
51	CPLUS_B_12H	Formate dehydrogenase H	ALNLLLBE_743668	76707	50%	1	1	1	0.01%	1.60%
52	CPLUS_B_12H	Respiratory nitrate reductase 1 beta chain	ALNLLLBE_218625	56752.5	76%	1	1	4	0.03%	5.80%
53	CPLUS_B_12H	hypothetical protein	ALNLLLBE_342544	70445.7	67%	2	2	5	0.04%	20.00%
54	CPLUS_B_12H	hypothetical protein	ALNLLLBE_331008	45626.9	77%	2	2	2	0.02%	7.70%

55	CPLUS_B_12H	hypothetical protein	**ALNLLLBE_68156**	93384.2	61%	1	1	2	0.02%	2.70%
56	CPLUS_B_12H	hypothetical protein	**ALNLLLBE_700955**	95628.6	97%	4	4	4	0.03%	4.70%
57	CPLUS_B_12H	hypothetical protein	**ALNLLLBE_675543**	67903.5	58%	1	1	1	0.01%	2.10%
58	CPLUS_B_12H	V-type ATP synthase beta chain	**ALNLLLBE_644750**	50785	32%	0	0	1	0.01%	1.70%
59	CPLUS_B_12H	V-type ATP synthase beta chain	**ALNLLLBE_54427**	50993	7%	0	0	1	0.01%	1.70%
60	CPLUS_B_12H	V-type ATP synthase alpha chain	**ALNLLLBE_644751**	63686.6	30%	0	0	1	0.01%	3.00%
61	CPLUS_B_12H	V-type ATP synthase alpha chain	**ALNLLLBE_54428**	63580	13%	0	0	1	0.01%	3.00%
62	CPLUS_B_12H	Outer membrane porin protein 32	**ALNLLLBE_296280**	33703.4	88%	2	2	4	0.03%	7.50%
63	CPLUS_B_12H	Outer membrane porin protein 32	**ALNLLLBE_104960**	38669.4	99%	2	2	2	0.02%	8.60%
64	CPLUS_B_12H	TonB-dependent receptor P3	**ALNLLLBE_506445**	115398.9	93%	3	3	3	0.02%	3.20%
65	CPLUS_B_12H	TonB-dependent receptor P3	**ALNLLLBE_75766**	118932.2	100%	7	7	8	0.06%	9.30%
66	CPLUS_B_12H	TonB-dependent receptor P26	**ALNLLLBE_76191**	115651	68%	2	2	2	0.02%	3.00%
67	CPLUS_B_12H	SusD-like protein P2	**ALNLLLBE_75767**	55877.7	93%	2	2	2	0.02%	3.10%
68	CPLUS_B_12H	Sialic acid-binding periplasmic protein SiaP	**ALNLLLBE_790261**	37481	57%	1	1	1	0.01%	4.50%
69	CPLUS_B_12H	Sialic acid-binding periplasmic protein SiaP	**ALNLLLBE_310190**	37510	57%	1	1	1	0.01%	4.50%
70	CPLUS_B_12H	Sialic acid-binding periplasmic protein SiaP	**ALNLLLBE_27688**	37511	57%	1	1	1	0.01%	4.50%
71	CPLUS_B_12H	Protein oar	**ALNLLLBE_612028**	105385	41%	1	1	1	0.01%	1.70%
72	CPLUS_B_12H	Periplasmic oligopeptide-binding protein	**ALNLLLBE_320014**	58261	64%	1	1	3	0.02%	5.20%
73	CPLUS_B_12H	Periplasmic oligopeptide-binding protein	**ALNLLLBE_326795**	58269	72%	2	2	4	0.03%	6.60%
74	CPLUS_B_12H	NAD-specific glutamate dehydrogenase	**ALNLLLBE_791579**	48866	35%	1	1	1	0.01%	4.00%
75	CPLUS_B_12H	Membrane lipoprotein TmpC	**ALNLLLBE_260417**	39364	43%	1	1	1	0.01%	4.10%
76	CPLUS_B_12H	60 kDa chaperonin 5	**ALNLLLBE_67393**	57748	24%	0	0	2	0.02%	4.00%
77	CPLUS_B_12H	60 kDa chaperonin 1	**ALNLLLBE_399492**	37451	48%	1	1	2	0.02%	6.60%
78	CPLUS_B_12H	60 kDa chaperonin	**ALNLLLBE_333915**	57697	42%	1	1	2	0.02%	3.40%

79	CPLUS_B_12H	60 kDa chaperonin	**ALNLLLBE_121083**	58153	68%	1	1	1	0.01%	3.10%
80	CPLUS_B_12H	3-hydroxylaminophenol mutase	**ALNLLLBE_142484**	52268	88%	1	1	1	0.01%	3.00%
81	CPLUS_B_12H	K(+)-insensitive pyrophosphate-energized proton pump	**ALNLLLBE_11910**	70658	93%	2	2	3	0.02%	5.20%
82	CPLUS_B_12H	K(+)-insensitive pyrophosphate-energized proton pump	**ALNLLLBE_389495**	57984	93%	2	2	3	0.02%	6.30%
83	CPLUS_B_12H	Cyanate hydratase	**ALNLLLBE_88274**	18009.5	88%	3	3	4	0.03%	17.00%
84	CPLUS_B_12H	Corrinoid/iron-sulfur protein large subunit	**ALNLLLBE_334543**	52053.7	58%	1	1	1	0.01%	1.90%
85	CPLUS_B_12H	Butyryl-CoA:acetate CoA-transferase	**ALNLLLBE_470572**	57360.9	95%	2	2	2	0.02%	5.20%
86	CPLUS_B_12H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_330937**	46094.9	44%	0	0	2	0.02%	4.00%
87	CPLUS_B_12H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_369805**	45940.6	44%	0	0	2	0.02%	4.00%
88	CPLUS_B_12H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_477611**	46315	10%	0	0	2	0.02%	4.00%
89	CPLUS_B_12H	Acetyl-coenzyme A synthetase	**ALNLLLBE_642943**	75311	36%	1	1	1	0.01%	1.80%
90	CPLUS_B_12H	ATP synthase subunit beta	**ALNLLLBE_343885**	51090	27%	1	1	1	0.01%	3.00%
91	CPLUS_B_12H	ATP synthase subunit alpha	**ALNLLLBE_413871**	53611.1	85%	1	1	2	0.02%	3.00%
92	CPLUS_B_12H	ALNLLLBE_614387-DECOY	**ALNLLLBE_614387-DECOY**		28%	1	1	1	0.01%	
1	C8_NB_0H	hypothetical protein	**ALNLLLBE_641604**	28724.7	100%	5	5	8	0.08%	28%
2	C8_NB_0H	hypothetical protein	**ALNLLLBE_642509**	63736.2	100%	8	8	11	0.11%	20%
3	C8_NB_0H	hypothetical protein	**ALNLLLBE_743383**	28982.8	14%	0	0	1	0.01%	4%
4	C8_NB_0H	hypothetical protein	**ALNLLLBE_641606**	45516.9	57%	1	1	1	0.01%	3%
5	C8_NB_0H	hypothetical protein	**ALNLLLBE_641603**	61392.1	100%	2	2	3	0.03%	4.50%
6	C8_NB_0H	Carbon monoxide dehydrogenase/acetyl-CoA synthase subunit alpha	**ALNLLLBE_334547**	51772.6	99%	3	3	4	0.04%	11.00%
7	C8_NB_0H	hypothetical protein	**ALNLLLBE_743321**	93204.8	32%	2	2	2	0.02%	2.70%
8	C8_NB_0H	hypothetical protein	**ALNLLLBE_641420**	24057.7	93%	2	2	3	0.03%	15%
9	C8_NB_0H	hypothetical protein	**ALNLLLBE_334549**	89357.7	19%	1	1	1	0.01%	2%

#	Sample	Description	Accession							
10	C8_NB_0H	Elongation factor Tu	ALNLLLBE_246749	43095.6	81%	1	1	2	0.02%	3.50%
11	C8_NB_0H	Respiratory nitrate reductase 2 beta chain	ALNLLLBE_173240	57135.2	18%	0	0	1	0.01%	2.40%
12	C8_NB_0H	Respiratory nitrate reductase 1 beta chain	ALNLLLBE_218625	56752	18%	0	0	1	0.01%	2.40%
13	C8_NB_0H	hypothetical protein	ALNLLLBE_641330	76715.7	83%	2	2	2	0.02%	3.00%
14	C8_NB_0H	hypothetical protein	ALNLLLBE_743382	62548	18%	0	0	1	0.01%	1.40%
15	C8_NB_0H	hypothetical protein	ALNLLLBE_643285	53571.2	54%	1	1	3	0.03%	9.50%
16	C8_NB_0H	hypothetical protein	ALNLLLBE_342549	41535	95%	2	2	2	0.02%	6.70%
17	C8_NB_0H	hypothetical protein	ALNLLLBE_438590	57521.6	62%	2	2	2	0.02%	4.60%
18	C8_NB_0H	hypothetical protein	ALNLLLBE_580468	78873	28%	1	1	2	0.02%	2.50%
19	C8_NB_0H	hypothetical protein	ALNLLLBE_259625	55559	83%	1	1	1	0.01%	3.30%
20	C8_NB_0H	hypothetical protein	ALNLLLBE_90799	42809.2	27%	1	1	1	0.01%	0.00%
21	C8_NB_0H	hypothetical protein	ALNLLLBE_342545	78943.6	74%	1	1	2	0.02%	7.20%
22	C8_NB_0H	hypothetical protein	ALNLLLBE_700955	95628.6	58%	2	2	2	0.02%	2.20%
23	C8_NB_0H	hypothetical protein	ALNLLLBE_474867	67684.4	39%	1	1	1	0.01%	0.96%
24	C8_NB_0H	hypothetical protein	ALNLLLBE_45809	28996	53%	1	1	1	0.01%	6.60%
25	C8_NB_0H	hypothetical protein	ALNLLLBE_761442	23483	8%	0	0	1	0.01%	5.50%
26	C8_NB_0H	hypothetical protein	ALNLLLBE_342544	70446	6%	0	0	1	0.01%	2.50%
27	C8_NB_0H	hypothetical protein	ALNLLLBE_105999	76403	66%	2	2	2	0.02%	3.60%
28	C8_NB_0H	hypothetical protein	ALNLLLBE_75826	83711	66%	2	2	2	0.02%	3.30%
29	C8_NB_0H	Respiratory nitrate reductase 1 beta chain	ALNLLLBE_584064	57747.9	41%	1	1	1	0.01%	2.40%
30	C8_NB_0H	Outer membrane porin protein 32	ALNLLLBE_296280	33703	36%	1	1	1	0.01%	3.00%
31	C8_NB_0H	K(+)-insensitive pyrophosphate-energized proton pump	ALNLLLBE_389495	57984	9%	0	0	1	0.01%	1.90%
32	C8_NB_0H	K(+)-insensitive pyrophosphate-energized proton pump	ALNLLLBE_11910	70658	9%	0	0	1	0.01%	1.60%
33	C8_NB_0H	ATP synthase subunit beta	ALNLLLBE_376596	42928	59%	1	1	1	0.01%	2.50%

34	C8_NB_0H	DNA-directed RNA polymerase subunit beta	**ALNLLLBE_454561**	154238	19%	1	1	1	0.01%	0.58%
35	C8_NB_0H	3-hydroxylaminophenol mutase	**ALNLLLBE_142484**	52267.8	71%	1	1	1	0.01%	3.00%
36	C8_NB_0H	60 kDa chaperonin 5	**ALNLLLBE_06125**	58496	71%	0	0	2	0.02%	3.50%
37	C8_NB_0H	60 kDa chaperonin 5	**ALNLLLBE_67393**	57748.3	94%	1	1	3	0.03%	5.50%
38	C8_NB_0H	60 kDa chaperonin 1	**ALNLLLBE_399492**	37451.1	37%	0	0	2	0.02%	5.80%
39	C8_NB_0H	60 kDa chaperonin	**ALNLLLBE_333915**	57697	5%	0	0	1	0.01%	1.30%
40	C8_NB_0H	60 kDa chaperonin	**ALNLLLBE_121083**	58153	16%	0	0	1	0.01%	1.30%
41	C8_NB_0H	Cyanate hydratase	**ALNLLLBE_88274**	18009.5	69%	2	2	2	0.02%	11%
42	C8_NB_0H	Asparagine synthetase [glutamine-hydrolyzing] 1	**ALNLLLBE_434877**	72205	45%	2	2	2	0.02%	3%
43	C8_NB_0H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_330937**	46094.9	26%	0	0	1	0.01%	4%
44	C8_NB_0H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_369805**	45940.6	26%	0	0	1	0.01%	4%
45	C8_NB_0H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_477611**	46315.1	6%	0	0	1	0.01%	4%
1	C8_NB_6H	hypothetical protein	**ALNLLLBE_641604**	28724.7	87%	2	2	2	0.03%	9%
2	C8_NB_6H	hypothetical protein	**ALNLLLBE_642509**	63736.2	100%	7	7	10	0.12%	15.00%
3	C8_NB_6H	hypothetical protein	**ALNLLLBE_641603**	61392.1	96%	3	3	3	0.04%	6%
4	C8_NB_6H	hypothetical protein	**ALNLLLBE_641420**	24058	93%	2	2	4	0.05%	15%
5	C8_NB_6H	hypothetical protein	**ALNLLLBE_743321**	93204.8	86%	2	2	2	0.03%	3%
6	C8_NB_6H	hypothetical protein	**ALNLLLBE_342549**	41535.4	98%	2	2	2	0.03%	7%
7	C8_NB_6H	Carbon monoxide dehydrogenase/acetyl-CoA synthase subunit alpha	**ALNLLLBE_334547**	51772.6	96%	1	1	1	0.01%	3.00%
8	C8_NB_6H	hypothetical protein	**ALNLLLBE_645567**	243258.6	95%	2	2	2	0.03%	1.30%
9	C8_NB_6H	hypothetical protein	**ALNLLLBE_641330**	76715.7	99%	4	4	4	0.05%	7.80%
10	C8_NB_6H	hypothetical protein	**ALNLLLBE_643285**	53571.2	6%	0	0	1	0.01%	1.40%
11	C8_NB_6H	hypothetical protein	**ALNLLLBE_580468**	78873	100%	5	6	8	0.10%	9.40%
12	C8_NB_6H	hypothetical protein	**ALNLLLBE_45809**	28995.9	43%	1	1	1	0.01%	6.60%

13	C8_NB_6H	hypothetical protein	**ALNLLLBE_438590**	57521.6	36%	1	1	1	0.01%	1.10%
14	C8_NB_6H	hypothetical protein	**ALNLLLBE_105999**	76402.6	85%	1	1	1	0.01%	2.20%
15	C8_NB_6H	hypothetical protein	**ALNLLLBE-75826**	83710.8	85%	1	1	1	0.01%	2.00%
16	C8_NB_6H	hypothetical protein	**ALNLLLBE_334511**	97269.8	62%	2	2	2	0.03%	2.90%
17	C8_NB_6H	hypothetical protein	**ALNLLLBE_259625**	55559	89%	2	2	2	0.03%	6.00%
18	C8_NB_6H	hypothetical protein	**ALNLLLBE_521398**	142014	69%	2	2	2	0.03%	2.10%
19	C8_NB_6H	hypothetical protein	**ALNLLLBE_669128**	31780	38%	1	1	2	0.03%	3.50%
20	C8_NB_6H	hypothetical protein	**ALNLLLBE_733678**	39382	49%	1	1	1	0.01%	4.40%
21	C8_NB_6H	hypothetical protein	**ALNLLLBE_487977**	154657	6%	0	0	1	0.01%	0.57%
22	C8_NB_6H	Elongation factor Tu	**ALNLLLBE_246749**	43095.6	92%	1	1	2	0.03%	3.50%
23	C8_NB_6H	Elongation factor Tu	**ALNLLLBE_203705**	43232.5	26%	1	1	1	0.01%	2.80%
24	C8_NB_6H	DNA-directed RNA polymerase subunit beta	**ALNLLLBE_454561**	154237.5	6%	0	0	1	0.01%	0.44%
25	C8_NB_6H	DNA-directed RNA polymerase subunit beta	**ALNLLLBE_277117**	89607	6%	0	0	1	0.01%	1.40%
26	C8_NB_6H	DNA-directed RNA polymerase subunit beta	**ALNLLLBE_687394**	158777	6%	0	0	1	0.01%	0.78%
27	C8_NB_6H	Respiratory nitrate reductase 2 beta chain	**ALNLLLBE_173240**	57135.2	55%	1	1	1	0.01%	3.20%
28	C8_NB_6H	Outer membrane porin protein 32	**ALNLLLBE_296280**	33703	29%	1	1	1	0.01%	3.00%
29	C8_NB_6H	K(+)-insensitive pyrophosphate-energized proton pump	**ALNLLLBE_11910**	70658	10%	0	0	1	0.01%	1.60%
30	C8_NB_6H	K(+)-insensitive pyrophosphate-energized proton pump	**ALNLLLBE_389495**	57984	10%	0	0	1	0.01%	1.90%
31	C8_NB_6H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_330937**	46094.9	84%	2	2	4	0.05%	10.00%
32	C8_NB_6H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_369805**	45940.6	84%	2	2	4	0.05%	10.00%
33	C8_NB_6H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_477611**	46315.1	5%	0	0	2	0.03%	4.00%
34	C8_NB_6H	ATP synthase subunit beta	**ALNLLLBE_376596**	42928	48%	1	1	1	0.01%	2.50%
35	C8_NB_6H	3-hydroxylaminophenol mutase	**ALNLLLBE_142484**	52267.8	78%	2	2	2	0.03%	5.10%
36	C8_NB_6H	60 kDa chepronin 5	**ALNLLLBE_67393**	57748.3	97%	1	1	3	0.04%	5.50%

37	C8_NB_6H	60 kDa chepronin 5	**ALNLLLBE_06125**	58496	77%	0	0	2	0.03%	3.50%
38	C8_NB_6H	60 kDa chepronin 1	**ALNLLLBE_399492**	37451.1	60%	1	1	3	0.04%	8.70%
39	C8_NB_6H	60 kDa chepronin	**ALNLLLBE_333915**	57697	8%	0	0	1	0.01%	1.30%
40	C8_NB_6H	60 kDa chepronin	**ALNLLLBE_131203**	49856	15%	0	0	2	0.03%	5.10%
41	C8_NB_6H	Cyanate hydratase	**ALNLLLBE_88274**	18009.5	81%	2	2	2	0.03%	11.00%
42	C8_NB_6H	ALNLLLBE_614387-DECOY	**ALNLLLBE_614387-DECOY**		28%	1	1	1	0.01%	
1	C8_NB_12H	hypothetical protein	**ALNLLLBE_641604**	28724.7	99%	4	4	5	0.04%	22%
2	C8_NB_12H	hypothetical protein	**ALNLLLBE_642509**	63736.2	100%	8	8	11	0.09%	20%
3	C8_NB_12H	hypothetical protein	**ALNLLLBE_641603**	61392.1	91%	2	2	2	0.02%	5%
4	C8_NB_12H	hypothetical protein	**ALNLLLBE_743383**	28983	70%	0	0	3	0.03%	9%
5	C8_NB_12H	hypothetical protein	**ALNLLLBE_641606**	45517	24%	1	1	1	0.01%	3%
6	C8_NB_12H	hypothetical protein	**ALNLLLBE_643285**	53571.2	6%	0	0	1	0.01%	1%
7	C8_NB_12H	hypothetical protein	**ALNLLLBE_641420**	24057.7	86%	1	1	2	0.02%	8%
8	C8_NB_12H	hypothetical protein	**ALNLLLBE_197203**	28469	39%	0	0	2	0.02%	5%
9	C8_NB_12H	hypothetical protein	**ALNLLLBE_334549**	89358	85%	3	3	4	0.03%	6%
10	C8_NB_12H	hypothetical protein	**ALNLLLBE_645567**	243259	70%	1	1	1	0.01%	1%
11	C8_NB_12H	hypothetical protein	**ALNLLLBE_54401**	89183	10%	0	0	1	0.01%	1%
12	C8_NB_12H	hypothetical protein	**ALNLLLBE_45809**	28996	81%	1	1	1	0.01%	7%
13	C8_NB_12H	hypothetical protein	**ALNLLLBE_219451**	23245	57%	2	2	3	0.03%	10%
14	C8_NB_12H	hypothetical protein	**ALNLLLBE_342545**	78944	14%	0	0	1	0.01%	2%
15	C8_NB_12H	hypothetical protein	**ALNLLLBE_90799**	42809	27%	1	1	1	0.01%	0%
16	C8_NB_12H	hypothetical protein	**ALNLLLBE_753512**	12463	95%	3	3	3	0.03%	24%
17	C8_NB_12H	hypothetical protein	**ALNLLLBE_764319**	96026	14%	0	0	1	0.01%	1%
18	C8_NB_12H	hypothetical protein	**ALNLLLBE_763559**	101797	10%	0	0	1	0.01%	1%
19	C8_NB_12H	Elongation factor Tu	**ALNLLLBE_246749**	43095.6	87%	2	2	3	0.03%	9%
20	C8_NB_12H	Elongation factor Tu	**ALNLLLBE_483213**	42932	97%	1	1	2	0.02%	7%

21	C8_NB_12H	Elongation factor Tu	**ALNLLLBE_716068**	42735	94%	2	2	2	0.02%	6%
22	C8_NB_12H	Elongation factor Tu	**ALNLLLBE_203705**	43233	6%	0	0	1	0.01%	2%
23	C8_NB_12H	DNA-directed RNA polymerase subunit beta	**ALNLLLBE_277117**	89607	65%	0	0	2	0.02%	3%
24	C8_NB_12H	DNA-directed RNA polymerase subunit beta	**ALNLLLBE_372517**	119264	6%	0	0	2	0.02%	2%
25	C8_NB_12H	DNA-directed RNA polymerase subunit beta	**ALNLLLBE_550114**	38019	18%	0	0	1	0.01%	3%
26	C8_NB_12H	Respiratory nitrate reductase 2 beta chain	**ALNLLLBE_173240**	57135.2	94%	1	1	2	0.02%	4.90%
27	C8_NB_12H	Respiratory nitrate reductase 1 beta chain	**ALNLLLBE_584064**	57748	98%	1	1	2	0.02%	4.90%
28	C8_NB_12H	Carbon monoxide dehydrogenase/acetyl-CoA synthase subunit alpha	**ALNLLLBE_334547**	51772.6	100%	2	2	3	0.03%	7.50%
29	C8_NB_12H	3-hydroxylaminophenol mutase	**ALNLLLBE_142484**	52267.8	87%	1	1	1	0.01%	3%
30	C8_NB_12H	hypothetical protein	**ALNLLLBE_342549**	41535.4	45%	1	1	1	0.01%	3.10%
31	C8_NB_12H	hypothetical protein	**ALNLLLBE_438590**	57521.6	37%	1	1	1	0.01%	1.10%
32	C8_NB_12H	hypothetical protein	**ALNLLLBE_474867**	67684.4	87%	2	2	2	0.02%	2.60%
33	C8_NB_12H	hypothetical protein	**ALNLLLBE_641342**	26714	97%	2	2	2	0.02%	10.00%
34	C8_NB_12H	hypothetical protein	**ALNLLLBE_105999**	76402.6	15%	1	1	1	0.01%	2.20%
35	C8_NB_12H	hypothetical protein	**ALNLLLBE_75826**	83710.8	15%	1	1	1	0.01%	2.00%
36	C8_NB_12H	Nitrogen regulatory protein P-II	**ALNLLLBE_426853**	14100	46%	1	1	1	0.01%	8.60%
37	C8_NB_12H	K(+)-insensitive pyrophosphate-energized proton pump	**ALNLLLBE_389495**	57984	16%	0	0	1	0.01%	1.90%
38	C8_NB_12H	K(+)-insensitive pyrophosphate-energized proton pump	**ALNLLLBE_11910**	70658	16%	0	0	1	0.01%	1.60%
39	C8_NB_12H	Glutamine synthetase	**ALNLLLBE_315415**	30249	21%	1	1	1	0.01%	5.20%
40	C8_NB_12H	60 kDa chaperonin 1	**ALNLLLBE_399492**	37451.1	57%	0	0	3	0.03%	7.50%
41	C8_NB_12H	60 kDa chaperonin 5	**ALNLLLBE_67393**	57748.3	100%	2	2	5	0.04%	9.70%
42	C8_NB_12H	60 kDa chaperonin 5	**ALNLLLBE_06125**	58496	94%	0	0	2	0.02%	3.50%
43	C8_NB_12H	60 kDa chaperonin 2	**ALNLLLBE_254983**	56897	9%	0	0	1	0.01%	1.30%

44	C8_NB_12H	60 kDa chaperonin	**ALNLLLBE_333915**	57697	9%	0	0	1	0.01%	1.30%
45	C8_NB_12H	60 kDa chaperonin	**ALNLLLBE_121083**	58153	5%	0	0	1	0.01%	1.30%
46	C8_NB_12H	60 kDa chaperonin	**ALNLLLBE_583384**	57865	6%	0	0	1	0.01%	1.30%
47	C8_NB_12H	Cold shock protein CspA	**ALNLLLBE_20439**	8612	90%	2	2	2	0.02%	20.00%
48	C8_NB_12H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_369805**	45941	67%	2	2	3	0.03%	9.70%
49	C8_NB_12H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_330937**	46095	67%	2	2	3	0.03%	9.60%
50	C8_NB_12H	ATP synthase subunit beta	**ALNLLLBE_376596**	42928	60%	1	1	1	0.01%	2.50%
51	C8_NB_12H	ATP synthase subunit beta	**ALNLLLBE_343885**	51090	15%	1	1	1	0.01%	3.00%
1	C8_NB_18H	hypothetical protein	**ALNLLLBE_641604**	28724.7	100%	5	5	6	0.06%	27%
2	C8_NB_18H	hypothetical protein	**ALNLLLBE_642509**	63736.2	85%	1	1	2	0.02%	3.20%
3	C8_NB_18H	hypothetical protein	**ALNLLLBE_641606**	45516.9	100%	6	6	6	0.06%	21%
4	C8_NB_18H	hypothetical protein	**ALNLLLBE_641603**	61392.1	100%	7	8	9	0.09%	12%
5	C8_NB_18H	hypothetical protein	**ALNLLLBE_438590**	57521.6	37%	1	1	1	0.01%	1%
6	C8_NB_18H	hypothetical protein	**ALNLLLBE_474867**	67684.4	96%	3	3	3	0.03%	4%
7	C8_NB_18H	Carbon monoxide dehydrogenase/acetyl-CoA synthase subunit alpha	**ALNLLLBE_334557**	51772.6	100%	4	4	5	0.05%	12%
8	C8_NB_18H	hypothetical protein	**ALNLLLBE_641420**	24057.7	63%	1	1	1	0.01%	4.90%
9	C8_NB_18H	hypothetical protein	**ALNLLLBE_743382**	62548	96%	2	2	2	0.02%	3.40%
10	C8_NB_18H	hypothetical protein	**ALNLLLBE_334549**	89357.7	78%	1	1	3	0.03%	3.20%
11	C8_NB_18H	hypothetical protein	**ALNLLLBE_743383**	28982.8	95%	1	1	5	0.05%	14.00%
12	C8_NB_18H	hypothetical protein	**ALNLLLBE_197203**	28469	92%	2	2	5	0.05%	13.00%
13	C8_NB_18H	hypothetical protein	**ALNLLLBE_641330**	76716	94%	2	2	2	0.02%	3.30%
14	C8_NB_18H	hypothetical protein	**ALNLLLBE_743321**	93205	38%	1	1	1	0.01%	1.60%
15	C8_NB_18H	hypothetical protein	**ALNLLLBE_54401**	89183	12%	0	0	2	0.02%	2.00%
16	C8_NB_18H	hypothetical protein	**ALNLLLBE_643285**	53571	5%	0	0	1	0.01%	1.40%
17	C8_NB_18H	hypothetical protein	**ALNLLLBE_743368**	25947.6	17%	1	1	1	0.01%	4.90%

#	Sample	Protein	Accession	Value	%				%	%
18	C8_NB_18H	hypothetical protein	ALNLLLBE_743370	25148.5	17%	1	1	1	0.01%	5.00%
19	C8_NB_18H	hypothetical protein	ALNLLLBE_743386	46186.3	77%	1	1	1	0.01%	1.60%
20	C8_NB_18H	hypothetical protein	ALNLLLBE_45809	28996	81%	2	2	3	0.03%	9.30%
21	C8_NB_18H	hypothetical protein	ALNLLLBE_669128	31779.5	25%	1	1	2	0.02%	3.50%
22	C8_NB_18H	hypothetical protein	ALNLLLBE_466048	14584	68%	1	1	1	0.01%	9.90%
23	C8_NB_18H	hypothetical protein	ALNLLLBE_90799	42809	21%	1	1	2	0.02%	0.00%
24	C8_NB_18H	hypothetical protein	ALNLLLBE_68156	93384	22%	0	0	1	0.01%	1.50%
25	C8_NB_18H	hypothetical protein	ALNLLLBE_653168	99054	31%	0	0	1	0.01%	1.40%
26	C8_NB_18H	hypothetical protein	ALNLLLBE_781328	95890	26%	1	1	1	0.01%	1.80%
27	C8_NB_18H	hypothetical protein	ALNLLLBE_781347	55467	26%	1	1	1	0.01%	3.10%
28	C8_NB_18H	V-type ATP synthase beta chain	ALNLLLBE_644750	50785	21%	1	1	1	0.01%	2.60%
29	C8_NB_18H	V-type ATP synthase alpha chain	ALNLLLBE_644751	63687	16%	1	1	1	0.01%	1.70%
30	C8_NB_18H	Elongation factor Tu	ALNLLLBE_246749	43095.6	90%	2	2	2	0.02%	6.60%
31	C8_NB_18H	Elongation factor Tu	ALNLLLBE_483213	42931.5	78%	0	0	2	0.02%	4.00%
32	C8_NB_18H	Elongation factor Tu	ALNLLLBE_716068	42735	99%	2	2	2	0.03%	6.40%
33	C8_NB_18H	Elongation factor Tu	ALNLLLBE_526058	43716	35%	1	1	1	0.01%	3.20%
34	C8_NB_18H	Respiratory nitrate reductase 2 beta chain	ALNLLLBE_173240	57135.2	77%	1	1	1	0.01%	3.20%
35	C8_NB_18H	Leu/Ille/Val-binding protein	ALNLLLBE_328832	41182.7	23%	1	1	1	0.01%	1.80%
36	C8_NB_18H	Leu/Ille/Val-binding protein	ALNLLLBE_343029	41248.7	23%	1	1	1	0.01%	1.80%
37	C8_NB_18H	Phthiodiolone/phenolphthiodiolone dimycocerosates ketoreductase	ALNLLLBE_743328	34701	66%	1	1	1	0.01%	3.00%
38	C8_NB_18H	Nitrogen regulatory protein P-II	ALNLLLBE_426853	14100	24%	1	1	1	0.01%	8.60%
39	C8_NB_18H	NAD-specific glutamate dehydrogenase	ALNLLLBE_791579	48866	47%	1	1	1	0.01%	4.00%
40	C8_NB_18H	Glutamine synthetase	ALNLLLBE_315415	30249	34%	1	1	1	0.01%	5.20%
41	C8_NB_18H	3-hydroxylaminophenol mutase	ALNLLLBE_142484	52267.8	87%	1	1	1	0.01%	3.00%
42	C8_NB_18H	ATP synthase subunit beta 1	ALNLLLBE_189018	50811	19%	1	1	1	0.01%	4.10%

43	C8_NB_18H	ATP synthase subunit beta	**ALNLLLBE_376596**	42928	85%	1	1	2	0.02%	6.00%
44	C8_NB_18H	ATP synthase subunit beta	**ALNLLLBE_343885**	51090	25%	1	1	1	0.01%	3.00%
45	C8_NB_18H	ATP synthase subunit beta	**ALNLLLBE_756712**	54705	19%	0	0	1	0.01%	2.80%
46	C8_NB_18H	ATP synthase subunit beta	**ALNLLLBE_441698**	34208.7	66%	1	1	2	0.02%	8.80%
47	C8_NB_18H	DNA-directed RNA-polymerase subunit beta	**ALNLLLBE_454561**	154237.5	6%	0	0	1	0.01%	0.44%
48	C8_NB_18H	DNA-directed RNA-polymerase subunit beta	**ALNLLLBE_277117**	89607	47%	0	0	2	0.02%	2.50%
49	C8_NB_18H	DNA-directed RNA-polymerase subunit beta	**ALNLLLBE_550114**	38019	15%	0	0	1	0.01%	2.60%
50	C8_NB_18H	Corrinoid/iron-sulfur protein large subunit	**ALNLLLBE_334543**	52054	56%	1	1	1	0.01%	1.70%
51	C8_NB_18H	Cold shock protein CspA	**ALNLLLBE_20439**	8611.9	70%	1	1	1	0.01%	8.90%
52	C8_NB_18H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_330937**	46094.9	71%	1	1	1	0.01%	3%
53	C8_NB_18H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_369805**	45940.6	71%	1	1	1	0.01%	3%
54	C8_NB_18H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_477611**	46315.1	15%	1	1	1	0.01%	2%
55	C8_NB_18H	60 kDa chaperonin 5	**ALNLLLBE_67393**	57748.3	98%	1	1	4	0.04%	6.50%
56	C8_NB_18H	60 kDa chaperonin 5	**ALNLLLBE_06125**	58496	78%	0	0	2	0.02%	3.50%
57	C8_NB_18H	60 kDa chaperonin 1	**ALNLLLBE_399492**	37451.1	60%	0	0	3	0.03%	7.50%
58	C8_NB_18H	60 kDa chaperonin 1	**ALNLLLBE_66011**	58576	11%	0	0	1	0.01%	1.30%
59	C8_NB_18H	60 kDa chaperonin 1	**ALNLLLBE_66021**	58521	11%	0	0	1	0.01%	1.30%
60	C8_NB_18H	60 kDa chaperonin	**ALNLLLBE_333915**	57697	9%	0	0	1	0.01%	1.30%
61	C8_NB_18H	60 kDa chaperonin	**ALNLLLBE_131203**	49856	13%	0	0	2	0.02%	5.10%
62	C8_NB_18H	60 kDa chaperonin	**ALNLLLBE_121083**	58153	5%	0	0	1	0.01%	1.30%
1	C8_NB_24H	hypothetical protein	**ALNLLLBE_641604**	28724.7	63%	0	0	1	0.01%	4%
2	C8_NB_24H	hypothetical protein	**ALNLLLBE_642509**	63736.2	60%	1	1	2	0.03%	3.20%
3	C8_NB_24H	hypothetical protein	**ALNLLLBE_438590**	57521.6	41%	1	1	1	0.01%	1%

4	C8_NB_24H	Carbon monoxide dehydrogenase/acetyl-CoA synthase subunit alpha	ALNLLLBE_334557	51772.6	79%	1	1	1	0.01%	3%
5	C8_NB_24H	hypothetical protein	ALNLLLBE_641420	24057.7	42%	1	1	1	0.01%	4.90%
6	C8_NB_24H	hypothetical protein	ALNLLLBE_743382	62548	48%	0	0	2	0.03%	1.40%
7	C8_NB_24H	hypothetical protein	ALNLLLBE_743383	28982.8	15%	0	0	1	0.01%	3.90%
8	C8_NB_24H	hypothetical protein	ALNLLLBE_197203	28469	57%	2	2	2	0.03%	7.50%
9	C8_NB_24H	hypothetical protein	ALNLLLBE_669128	31779.5	17%	1	1	1	0.01%	3.50%
10	C8_NB_24H	hypothetical protein	ALNLLLBE_90799	42809	48%	1	1	1	0.01%	0.00%
11	C8_NB_24H	hypothetical protein	ALNLLLBE_753512	12463	18%	1	1	1	0.01%	6.70%
12	C8_NB_24H	Elongation factor Tu	ALNLLLBE_246749	43095.6	27%	1	1	1	0.01%	3.00%
13	C8_NB_24H	Respiratory nitrate reductase 2 beta chain	ALNLLLBE_173240	57135.2	58%	1	1	2	0.03%	5.50%
14	C8_NB_24H	Respiratory nitrate reductase 1 beta chain	ALNLLLBE_218625	56752	7%	0	0	1	0.01%	2.40%
15	C8_NB_24H	Glutamine synthetase	ALNLLLBE_315415	30249	44%	1	1	1	0.01%	5.20%
16	C8_NB_24H	3-hydroxylaminophenol mutase	ALNLLLBE_142484	52267.8	67%	1	1	1	0.01%	3.00%
17	C8_NB_24H	DNA-directed RNA-polymerase subunit beta	ALNLLLBE_277117	89607	30%	0	0	1	0.01%	1.10%
18	C8_NB_24H	DNA-directed RNA-polymerase subunit beta	ALNLLLBE_550114	38019	16%	0	0	1	0.01%	2.60%
19	C8_NB_24H	Asparagine synthetase [glutamine-hydrolyzing] 1	ALNLLLBE_434877	72205	20%	1	1	1	0.01%	1.40%
20	C8_NB_24H	60 kDa chaperonin 5	ALNLLLBE_67393	57748.3	54%	0	0	2	0.03%	3.70%
21	C8_NB_24H	60 kDa chaperonin 5	ALNLLLBE_06125	58496	61%	0	0	2	0.03%	3.50%
22	C8_NB_24H	60 kDa chaperonin 1	ALNLLLBE_399492	37451.1	43%	0	0	2	0.03%	5.80%
23	C8_NB_24H	60 kDa chaperonin	ALNLLLBE_333915	57697	7%	0	0	1	0.01%	1.30%
1	CNEG_NB_12H	hypothetical protein	ALNLLLBE_641604	28724.7	94%	1	1	2	0.02%	10%
2	CNEG_NB_12H	hypothetical protein	ALNLLLBE_642509	63736.2	96%	3	3	4	0.05%	7.60%
3	CNEG_NB_12H	hypothetical protein	ALNLLLBE_641603	61392.1	100%	2	2	3	0.04%	5%
4	CNEG_NB_12H	hypothetical protein	ALNLLLBE_45809	28995.9	98%	2	2	2	0.02%	9%

#		Protein	Accession							
5	CNEG_NB_12H	hypothetical protein	ALNLLLBE_438590	57521.6	36%	1	1	1	0.01%	1%
6	CNEG_NB_12H	hypothetical protein	ALNLLLBE_474867	67684.4	43%	1	1	1	0.01%	2%
7	CNEG_NB_12H	Carbon monoxide dehydrogenase/acetyl-CoA synthase subunit alpha	ALNLLLBE_334557	51772.6	20%	1	1	1	0.01%	3%
8	CNEG_NB_12H	hypothetical protein	ALNLLLBE_743382	62548	15%	0	0	1	0.01%	1.40%
9	CNEG_NB_12H	hypothetical protein	ALNLLLBE_743383	28982.8	22%	0	0	1	0.01%	3.90%
10	CNEG_NB_12H	hypothetical protein	ALNLLLBE_197203	28469	90%	2	2	2	0.02%	7.50%
11	CNEG_NB_12H	hypothetical protein	ALNLLLBE_743386	46186.3	18%	1	1	1	0.01%	1.60%
12	CNEG_NB_12H	hypothetical protein	ALNLLLBE_90799	42809.2	64%	2	2	2	0.02%	0.00%
13	CNEG_NB_12H	hypothetical protein	ALNLLLBE_669128	31779.5	14%	1	1	1	0.01%	3.50%
14	CNEG_NB_12H	hypothetical protein	ALNLLLBE_259625	55559	97%	2	2	3	0.04%	6.00%
15	CNEG_NB_12H	hypothetical protein	ALNLLLBE_641342	26714	20%	1	1	1	0.01%	2.90%
16	CNEG_NB_12H	hypothetical protein	ALNLLLBE_219451	23245	41%	1	1	1	0.01%	6.00%
17	CNEG_NB_12H	hypothetical protein	ALNLLLBE_781328	95890	69%	1	1	1	0.01%	1.80%
18	CNEG_NB_12H	hypothetical protein	ALNLLLBE_781347	55467	69%	1	1	1	0.01%	3.10%
19	CNEG_NB_12H	Nitrogen regulatory protein P-II	ALNLLLBE_426853	14100	37%	1	1	1	0.01%	8.60%
20	CNEG_NB_12H	Elongation factor Tu	ALNLLLBE_246749	43095.6	99%	2	2	2	0.02%	6.60%
21	CNEG_NB_12H	Elongation factor Tu	ALNLLLBE_483213	42931.5	42%	0	0	1	0.01%	4.00%
22	CNEG_NB_12H	Elongation factor Tu	ALNLLLBE_203705	43233	29%	1	1	2	0.01%	2.80%
23	CNEG_NB_12H	Elongation factor Tu	ALNLLLBE_716068	42735	38%	1	1	1	0.01%	2.30%
24	CNEG_NB_12H	3-hydroxylaminophenol mutase	ALNLLLBE_142484	52267.8	84%	1	1	2	0.02%	3.00%
25	CNEG_NB_12H	ATP synthase subunit beta 1	ALNLLLBE_189018	50810.7	39%	1	1	1	0.01%	4.10%
26	CNEG_NB_12H	ATP synthase subunit beta	ALNLLLBE_376596	42928	70%	1	1	2	0.02%	5.20%
27	CNEG_NB_12H	ATP synthase subunit beta	ALNLLLBE_756712	53611	20%	0	0	1	0.01%	2.20%
28	CNEG_NB_12H	DNA-directed RNA-polymerase subunit beta	ALNLLLBE_454561	154238	6%	0	0	1	0.01%	0.44%
29	CNEG_NB_12H	60 kDa chaperonin 5	ALNLLLBE_67393	57748.3	98%	1	1	2	0.02%	3.10%

30	CNEG_NB_12H	60 kDa chaperonin 5	ALNLLLBE_06125	58496	84%	0	0	2	0.02%	3.50%
31	CNEG_NB_12H	60 kDa chaperonin 1	ALNLLLBE_399492	37451.1	9%	0	0	1	0.01%	2.00%
32	CNEG_NB_12H	60 kDa chaperonin	ALNLLLBE_333915	57697	8%	0	0	1	0.01%	1.30%
33	CNEG_NB_12H	60 kDa chaperonin	ALNLLLBE_747290	57589	64%	1	1	2	0.02%	3.70%
34	CNEG_NB_12H	60 kDa chaperonin	ALNLLLBE_121083	58153	78%	1	1	2	0.02%	4.40%
1	CPLUS_NB_12H	putative protein	ALNLLLBE_110201	35565	87%	2	2	2	0.02%	6%
2	CPLUS_NB_12H	putative protein	ALNLLLBE_162024	35551	87%	2	2	2	0.02%	6%
3	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_641604	28724.7	100%	11	14	18	0.18%	51%
4	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_642509	63736.2	87%	2	2	2	0.02%	5%
5	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_641606	45516.9	100%	10	10	12	0.12%	29%
6	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_641603	61392.1	100%	8	9	13	0.13%	16%
7	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_743321	93204.8	100%	8	8	8	0.08%	11%
8	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_641420	24057.7	87%	1	1	1	0.01%	5%
9	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_743382	62548	100%	5	5	6	0.06%	9.70%
10	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_334549	89357.7	100%	7	8	16	0.16%	20.00%
11	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_743383	28982.8	100%	4	4	5	0.05%	23.00%
12	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_645567	243258.6	94%	3	3	3	0.03%	1.60%
13	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_641330	76715.7	100%	7	7	8	0.08%	16.00%
14	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_197203	28469	74%	1	1	2	0.02%	9.10%
15	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_54401	89183	99%	2	2	10	0.10%	11.00%
16	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_259625	55559	99%	1	1	1	0.01%	3.30%
17	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_521398	142014	100%	5	5	5	0.05%	4.80%
18	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_474867	67684	30%	1	1	1	0.01%	0.96%
19	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_743386	46186	100%	5	6	6	0.06%	12.00%
20	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_580468	78873	98%	5	5	5	0.05%	8.70%
21	CPLUS_NB_12H	hypothetical protein	ALNLLLBE_743370	25148	36%	1	1	1	0.01%	5.00%

22	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_743368**	25948	36%	1	1	1	0.01%	4.90%
23	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_466046**	65802	78%	1	1	1	0.01%	1.70%
24	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_669128**	31780	33%	1	1	1	0.01%	3.50%
25	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_219451**	23245	66%	1	1	1	0.01%	4.70%
26	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_466048**	14584	97%	3	3	3	0.03%	41.00%
27	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_733678**	39382	88%	2	2	2	0.02%	8.80%
28	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_700955**	95629	99%	6	6	6	0.06%	7.60%
29	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_624953**	5654	26%	1	1	1	0.01%	24.00%
30	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_366056**	5654	26%	1	1	1	0.01%	24.00%
31	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_90799**	42809	36%	1	1	1	0.01%	0.00%
32	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_105999**	76403	28%	1	1	1	0.01%	2.20%
33	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_75826**	83711	28%	1	1	1	0.01%	2.00%
34	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_68156**	93384	8%	0	0	1	0.01%	1.50%
35	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_653168**	99054	24%	0	0	1	0.01%	1.40%
36	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_90712**	64996	59%	2	2	2	0.02%	4.70%
37	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_744241**	69352	49%	1	1	1	0.01%	2.20%
38	CPLUS_NB_12H	hypothetical protein	**ALNLLLBE_334544**	48476	69%	2	2	2	0.02%	3.30%
39	CPLUS_NB_12H	Elongation factor Tu	**ALNLLLBE_246749**	43095.6	82%	2	3	3	0.03%	7%
40	CPLUS_NB_12H	Elongation factor Tu	**ALNLLLBE_203705**	43233	17%	0	0	1	0.01%	3%
41	CPLUS_NB_12H	Carbon monoxide dehydrogenase/acetyl-CoA synthase subunit alpha	**ALNLLLBE_334547**	51772.6	100%	6	6	10	0.10%	21.00%
42	CPLUS_NB_12H	Carbon monoxide dehydrogenase/acetyl-CoA synthase subunit alpha	**ALNLLLBE_54403**	51728	29%	0	0	3	0.03%	6.40%
43	CPLUS_NB_12H	Respiratory nitrate reductase 2 beta chain	**ALNLLLBE_173240**	57135	68%	1	1	1	0.01%	3.20%
44	CPLUS_NB_12H	Glycerol kinase	**ALNLLLBE_232721**	55014.5	97%	3	3	3	0.03%	6.20%
45	CPLUS_NB_12H	Formate dehydrogenase H	**ALNLLLBE_743668**	76707	51%	1	1	1	0.01%	1.60%

46	CPLUS_NB_12H	V-type ATP synthase beta chain	**ALNLLLBE_644750**	50785	82%	0	0	2	0.02%	6.10%
47	CPLUS_NB_12H	V-type ATP synthase beta chain	**ALNLLLBE_54427**	50993	46%	1	1	3	0.03%	8.60%
48	CPLUS_NB_12H	V-type ATP synthase alpha chain	**ALNLLLBE_644751**	63686.6	99%	2	2	6	0.06%	12.00%
49	CPLUS_NB_12H	V-type ATP synthase alpha chain	**ALNLLLBE_54428**	63580	86%	1	1	5	0.05%	9.40%
50	CPLUS_NB_12H	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1	**ALNLLLBE_677586**	28755	69%	2	2	2	0.02%	7.10%
51	CPLUS_NB_12H	Outer membrane porin protein 32	**ALNLLLBE_296280**	33703.4	60%	1	1	1	0.01%	3.00%
52	CPLUS_NB_12H	Sialic acid-binding periplasmic protein SiaP	**ALNLLLBE_790261**	37481	95%	3	3	4	0.04%	12.00%
53	CPLUS_NB_12H	Sialic acid-binding periplasmic protein SiaP	**ALNLLLBE_310190**	37510	95%	3	3	4	0.04%	12.00%
54	CPLUS_NB_12H	Sialic acid-binding periplasmic protein SiaP	**ALNLLLBE_27688**	37511	95%	3	3	4	0.04%	12.00%
55	CPLUS_NB_12H	Phthiodiolone/phenolphthiodiolone dimycocerosates ketoreductase	**ALNLLLBE_743328**	34701	79%	2	2	2	0.02%	6.70%
56	CPLUS_NB_12H	Periplasmic [NiFeSe] hydrogenase large subunit	**ALNLLLBE_743277**	50212	26%	0	0	1	0.01%	2.40%
57	CPLUS_NB_12H	Periplasmic [NiFeSe] hydrogenase large subunit	**ALNLLLBE_465720**	50822	62%	1	1	2	0.02%	4.40%
58	CPLUS_NB_12H	Outer membrane protein A	**ALNLLLBE_90397**	49632	16%	1	1	1	0.01%	5.60%
59	CPLUS_NB_12H	NAD-specific glutamate dehydrogenase	**ALNLLLBE_791579**	48866	21%	1	1	1	0.01%	4.00%
60	CPLUS_NB_12H	Membrane lipoprotein TmpC	**ALNLLLBE_260417**	39364	61%	2	2	2	0.02%	7.90%
61	CPLUS_NB_12H	DNA-directed RNA polymerase subunit beta'	**ALNLLLBE_277117**	89607	19%	1	1	1	0.01%	1.70%
62	CPLUS_NB_12H	60 kDa chaperonin 5	**ALNLLLBE_67393**	57748	54%	1	1	2	0.02%	4.20%
63	CPLUS_NB_12H	60 kDa chaperonin 5	**ALNLLLBE_06125**	58496	61%	0	0	1	0.01%	2.20%
64	CPLUS_NB_12H	60 kDa chaperonin 1	**ALNLLLBE_399492**	37451	16%	0	0	1	0.01%	3.80%
65	CPLUS_NB_12H	60 kDa chaperonin	**ALNLLLBE_131203**	49856	11%	0	0	1	0.01%	3.60%
66	CPLUS_NB_12H	3-hydroxylaminophenol mutase	**ALNLLLBE_142484**	52268	89%	2	2	2	0.02%	5.10%
67	CPLUS_NB_12H	Cyanate hydratase	**ALNLLLBE_88274**	18009.5	89%	2	2	2	0.02%	12.00%
68	CPLUS_NB_12H	Corrinoid/iron-sulfur protein large subunit	**ALNLLLBE_334543**	52053.7	98%	3	3	3	0.03%	6.40%

69	CPLUS_NB_12H	Butyryl-CoA:acetate CoA-transferase	**ALNLLLBE_470572**	57360.9	51%	2	2	2	0.02%	5.20%
70	CPLUS_NB_12H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_330937**	46094.9	15%	0	0	1	0.01%	4.00%
71	CPLUS_NB_12H	Aliphatic amidase expression-regulating protein	**ALNLLLBE_369805**	45940.6	15%	0	0	1	0.01%	4.00%
72	CPLUS_NB_12H	Acetyl-coenzyme A synthetase	**ALNLLLBE_642943**	75311	98%	2	2	2	0.03%	5.80%
73	CPLUS_NB_12H	Acetyl-coenzyme A synthetase	**ALNLLLBE_642942**	75029	99%	4	4	4	0.04%	6.40%
74	CPLUS_NB_12H	ATP synthase subunit beta 1	**ALNLLLBE_189018**	50811	32%	1	1	1	0.01%	2.60%
75	CPLUS_NB_12H	ATP synthase subunit beta	**ALNLLLBE_376596**	42928	62%	1	1	1	0.01%	2.50%
76	CPLUS_NB_12H	ATP synthase subunit beta	**ALNLLLBE_343885**	51090	27%	1	1	1	0.01%	3.00%
77	CPLUS_NB_12H	50S ribosomal protein L1	**ALNLLLBE_246746**	23966	22%	1	1	1	0.01%	5.60%
78	CPLUS_NB_12H	ALNLLLBE_45281-DECOY	**ALNLLLBE_45281-DECOY**		17%	1	1	1	0.01%	%
79	CPLUS_NB_12H	ALNLLLBE_04111-DECOY	**ALNLLLBE_04111-DECOY**		44%	1	1	2	0.02%	

Appendix I

Table 4. The fractions of proteins in the abundant genera that were detected using metaproteomic data generated in the course of octanoic acid degradation.

genus	kraken id	gene	KEGG_ko	C8_B_0H	C8_N B_0H	C8_B_6 H	C8_N B_6H	C8_B_12H	C8_N B_12 H	C8_B_18H	C8_N B_18 H	C8_B_24H	C8_N B_24 H
NA	NA	NA	NA	1	2	2	2	1	2	2	2	2	2
Ruminococcus	Firm-04	NA	NA		2	1	1	2	1				
NA	NA	NA	NA	1	1		1	1	1	1		1	
Syntrophaceae bacterium	UBA8904	groL	ko:K04077		1			1	1	1	1	2	
Agathobacter	NA	groL	ko:K04077	2			2				2		
Verrucomicrobia subdivision 3	UBA6082	glnA	ko:K0915,k:K20712	1	1	2	2	1	1	1	1	1	1
Microbacterium	NA	narH	ko:K00371	5	1	3	1	5	2	5	1	1	2
Methanolinea	NA	mcr G	ko:K00402	4		3		6	2	10	5	8	2
Streptomyces	NA	tuf	ko:K02358			1	1		1	3		1	
Rikenellaceae bacterium	DMER64	NA	ko:K03704	3			1		2	1	1	1	
Acidovorax	NA	narH	ko:K00371	2	1	1		3		3			1
Smithella	NA	NA	ko:K12980	1					3			1	
Leptolinea	UBA4782	glpK	ko:K00864					1					
NA	NA	NA	NA					1					
Thermoleophilia bacterium	UBA2241	tuf	ko:K02358	4	2	1	2	3	3	3	2	3	1
NA	NA	NA	NA						1				
Bacteroidales bacterium	UBA5429	NA	ko:K21571	2	1		2	1		2		2	
Rectinema	UBA8932	siaP	ko:K21395					2					
Syntrophaceae bacterium	UBA2192	rpoC	ko:K03046,ko:K13797	1			1		2		2	1	1
NA	NA	NA	NA	1	1		1	1		1		1	

Rectinema	UBA8932	siaP	ko:K21395					2					
Verrucomicrobia subdivision 3	UBA6082	glnA	ko:K01915,ko:K20712						1	2	1		1
NA	NA	NA	NA							2	1	1	
Clostridium	NA	urtA	ko:K01999,ko:K11959	5	1	1	4	1	3	3	1	3	
NA	NA	NA	NA	1	1	1	1		1	1	1	2	1
Methanothrix	NA	NA	NA	3		1	2						
Methanothrix	NA	cdhC	ko:K00192,ko:K00193	9	4	4	1	4	3				
Methanothrix	NA	cdhA	ko:K00192	6	1	4		2	4	6	1	4	
Methanothrix	NA	NA	NA							7	4	5	1
Methanothrix	NA	NA	NA	1	1				2	1			
Methanothrix	NA	NA	ko:K01997	1	2				3	1	1		
Methanothrix	NA	NA	NA	3	2	3	2	1	1	1			
NA	NA	NA	NA							2	1	1	
NA	NA	NA	NA	2					1	1	1	1	
NA	NA	NA	NA	1					1		1	2	
Polaromonas	NA	urtA	ko:K01999,ko:K11959	5	1	1	4	1	3	3	1	3	
NA	NA	NA	NA						2				
NA	NA	NA	NA		1		1	2	1	2	2		
NA	NA	NA	NA	1	1		1	1	1	1		1	
Anaerolineaceae bacterium	49-20	glpK	ko:K00864						2				
Streptomyces	NA	groL	ko:K04077	1	2	2	3	1	3	3	3	3	2
Syntrophaceae bacterium	UBA8904	NA	NA						2				
NA	NA	NA	NA						2				
NA	NA	NA	NA	1					1		1	2	
Syntrophaceae bacterium	UBA2192	NA	ko:K01953		2								1
Planctomycetes bacterium	UBA8898	mce2D	ko:K02067	1	2	1	1		1	1	1	1	1

Syntrophorhabdus	NA	atpD	ko:K02112					1		1	2		
Syntrophaceae bacterium	UBA8904	rpoB	ko:K03043,ko:K13797		1	1	1			2	1	1	
NA	NA	NA	NA		1	1	1		1	3	3	2	
NA	NA	NA	NA					1					
Methanospirillum	UBA288	mcrG	ko:K00402					1		1	1		
NA	NA	NA	NA	1	1	2		1	2	1	3	1	
Syntrophaceae bacterium	UBA2192	urtA	ko:K01999,ko:K11959	1	1	1	2			3	1		
Thermoleophilia bacterium	UBA2241	NA	NA	1				1	2	2	2		
Methanothrix	NA	NA	NA				1						
Bacteroidales bacterium	UBA6192	rpoB	ko:K03043,ko:K13797									2	
Verrucomicrobia subdivision 3	UBA6082	NA	NA				2						
Streptomyces	NA	tuf	ko:K02358								1		
Methanothrix	NA	cdhA	ko:K00192			2		1	1	2	2	3	
NA	NA	NA	NA					1					
Methanothrix	NA	atpA	ko:K02117					1		1			
NA	NA	NA	NA						1		1	1	1
Bacteroidales bacterium	UBA5429	rpoB	ko:K03043,ko:K13797									2	
Methanothrix	NA	NA	NA	3	2	2	8						
NA	NA	NA	NA	1					1			2	
Microbacterium	NA	narH	ko:K00371	3	1	4		3	2	3		1	
Xanthomonadales bacterium	RPQJ01	tuf	ko:K02358							2			
NA	NA	NA	NA	1			1					1	
NA	NA	NA	NA	1					1		2		
Methanothrix	NA	NA	NA	7	2	1	4	5		1	2		
Methanothrix	NA	NA	NA			3		1	2	1			
Methanothrix	NA	NA	NA	4	3	2	4	6	2	1	1	1	1

Methanothrix	NA	mcrA	ko:K00399	6	3	6	3	6	2	3	8	4	
Methanothrix	NA	mcrG	ko:K00402	14	8	12	2	16	5	9	6	5	1
Methanothrix	NA	mcrB	ko:K00401	4	1	1		6	1	2	6	1	
Methanothrix	NA	NA	ko:K03006,ko:K13735	18	11	5	10	15	11	3	2	4	2
NA	NA	NA	NA	3		1		1					
Methanothrix	NA	NA	ko:K01895					1					
Methanothrix	NA	NA	ko:K03006,ko:K13735	4	3	2	1	1	1		1		
Methanothrix	NA	atpB	ko:K02118					1			1		
NA	NA	NA	NA					2					
Prolixibacteraceae bacterium	UBA1413	NA	NA	2		3	2	5	1				
Candidatus Hydrogenedentes	UBA2224	fla	ko:K02406									1	
NA	NA	NA	NA									1	
NA	NA	NA	NA									1	
Tahibacter	NA	NA	NA				2	2		2	2	4	1
Streptomyces	NA	groL	ko:K04077	3	3	5	3	1	5	5	4	3	2
Firmicutes bacterium	UBA3907	NA	NA					2					
Petrotoga	UBA5851	fla	ko:K02406									1	
NA	NA	NA	NA				1						
Firmicutes bacterium	UBA4882	NA	ko:K02035	1	2								
NA	NA	NA	NA						2	2	2	1	
Synergistaceae bacterium	58-81	NA	ko:K01999				1						
Methanolinea	NA	NA	NA	4	2	3	2	2		1	1	1	
NA	NA	NA	NA							1	1		
Methanolinea	NA	mtrA	ko:K00577					1		3	1	2	
Methanolinea	NA	mtrA	ko:K00577					1		3	1	2	

223

Taxon	Code	Gene	KO										
Methanolinea	NA	mcrA	ko:K00399		1	1				2	2	3	2
Methanolinea	NA	mcrG	ko:K00402	4	1	3		8	3	9	5	8	1
Methanolinea	NA	mcrB	ko:K00401					1		2	1	1	
Methanolinea	NA	fdhA	ko:K00123					1		1			
NA	NA	NA	NA						3	1			1
NA	NA	NA	NA					2					
NA	NA	NA	NA					2		1	1		
Ruminococcus	Firm-04	NA	NA		2	1	1	2	1				
NA	NA	NA	NA		1	1		2		2			
NA	NA	NA	NA						1				
NA	NA	NA	NA						1				
Thermoleophilia bacterium	UBA2241	mauB	ko:K01224,ko:K13372,ko:K15229,ko:K17285								1		
NA	NA	NA	NA						1				
Spirochaetaceae bacterium	UBA8932	siaP	ko:K21395					2					
Bacteroidales bacterium	UBA5429	gdh	ko:K00262					1		2	1		
Flexilinea	NA	cynS	ko:K01725	2	2	1	2	2					
NA	NA	NA	NA	1		1		1					
NA	NA	NA	NA	1	1	2		1	1		2		1
NA	NA	NA	NA					2					
Rikenellaceae bacterium	DMER64	atpD	ko:K02112	3							1		

Abbreviations

$^\circ$C	degrees celsius
ABC	ATP-Binding Cassette
ABCS	Affinity-Based Cell Sorting
ACO	Acyl-CoA Oxidase
AD	Anaerobic Digestion
ADBA	The Anaerobic Digestion and Bioresources Association
AHA	L-Azidohomoalanine
ASV	Amplicon Sequence Variant
ATP	Adenosine triphosphate
BC	Before Christ
BD	Exponential growth before the diauxic shift
BLAST	Basic Local Alignment Search Tool
BONCAT	Bioorthogonal non-canonical amino acid tagging
BRENDA	Braunschweig ENzyme DAtabase
bp	Base pair
BtG	Biomethane to the grid
BrdU	Bromodeoxyuridine
Cels.	Cellulosome
CH_4	Methane
C_2	Acetic acid
C_3	Propionic acid
C_4	Butyric acid
iC_4	Iso-butyric acid
C_5	Pentanoic acid
iC_5	Iso-pentanoic acid
C_6	Hexanoic acid
iC_6	Iso-hexanoic acid
C_7	Heptanoic acid
C_8	Octanoic acid
C_9	Nonanoic acid

C_{10}	Decanoic acid
Cm	Chloramphenicol
CO_2	Carbon dioxide
CoA	Coenzyme A
COD	Chemical Oxygen Demand
CoEMS	The York Centre of Excellence in Mass Spectrometry
COGs	Clusters of Orthologous Groups of proteins
C/N	Carbon to Nitrogen ratio
CQ	Quantitation cycle
CsCl	Caesium chloride
d	Day
DAPI	4′,6-diamidino-2-phenylindole
DGGE	Denaturing gradient gel electrophoresis
DIET	Direct interspecies electron transfer
DI/DII/DIII	During the diauxic shift
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded Deoxyribonucleic acid
DTT	Dithiothreitol
ECFA	Even-chain fatty acid
ECH	Enoyl-CoA hydratase
EDTA	Ethylenediaminetetraacetic acid
EJ	Exajoule
FA	Fatty acid
FAD^+	Oxidised flavin adenine dinucleotide
$FADH_2$	Reduced flavin adenine dinucleotide
FDH	Formate dehydrogenase
FDR	False Discovery Rate
FOG	Fat, oil, and grease
g	Gram
GC	Gas chromatography
GC-FID	Gas chromatography with flame-ionisation detection
GHG	Greenhouse gasses

GO	Gene Ontology
GTDB	The Genome Taxonomy Database
h	Hour
H_2	Hydrogen
HPG	L-homoproparglycine
HRT	Hydraulic retention time
IFT	Interspecies formate transfer
IHT	Interspecies hydrogen transfer
IMG/M	Integrated Microbial Genomes & Microbiomes
ISR	Inoculum-substrate ratio
k	Rate of substrate utilisation
Kan	Kanamycin
kbp	Kilobase pairs
KEGG	Kyoto Encyclopedia of Genes and Genomes
Kg	Kilogram
K_h	Hydrolysis rate constant
K_s	Half-saturation coefficient or substrate affinity
L	Litre
LB	Luria-Bertani
LCFA	Long-chain fatty acid
LC-MS/MS	Liquid chromatography-tandem-mass spectrometry
LEfSe	Linear discriminant analysis Effect Size
m^3	Cubic metre
μ_{max}	Specific growth rate
μm	Micrometre
μm^2	Square micrometre
μg	Microgram
μL	Microlitre
μM	Micromolar
M	Molar
mL	Millilitre
MAGs	Metagenome-assembled genomes

MCFA	Medium-chain fatty acid
Met	L-Methionine
MetRS	Methionyl-tRNA synthetase
mg	Milligram
MG-RAST	Metagenomics Rapid Annotation using Subsystem Technology
mol	Mole
mM	Millimolar
MOPS	Morpholinepropanesulfonic acid
$MtCO_2$	Metric tons of carbon dioxide equivalent
N_2	Nitrogen
NAD^+	Oxidised Nicotinamide adenine dinucleotide
$NADH_2$	Reduced Nicotinamide adenine dinucleotide
NaOH	Sodium hydroxide
NaCl	Sodium Chloride
NCBI	National Center for Biotechnology Information
ND	Not detected or not determined
NEB	New England Biolabs
NGS	Next-generation sequencing
nm	Nanometre
nL/min	Nanolitre per minute
ns	Not significant
OCFA	Odd-chain fatty acid
OD570	The optical density at 570 nm
OD600	The optical density at 600 nm
OLR	Organic loading rate
ONT	Oxford Nanopore Technologies'
ORF	Open Reading Frame
OTU	Operational taxonomic unit
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PFA	Paraformaldehyde

PICRUSt	Phylogenetic Investigation of Communities by Reconstruction of Unobserved States
pKa	Acid dissociation constant
QIIME2	Quantitative Insights Into Microbial Ecology version 2
qPCR	Quantitative polymerase chain reaction
RDP	Ribosomal database project
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
16S rRNA	16S ribosomal RNA
RT	Room temperature
SCFA	Short-chain fatty acid
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SIP	Stable-isotope probing
SMRT	Single Molecule Real-Time
SSU	The small subunit of ribosomes
Sus	Starch utilization system
TBE	Tris/Borate/EDTA buffer
TE	Tris and EDTA buffer
Temp.	Temperature
THP	Thermal hydrolysis process
T-RFLP	Terminal restriction fragment length polymorphism
UK	The United Kingdom
UV	Ultraviolet
v/v	Volume per volume
VBNC	Viable but non-culturable
VFA	Volatile fatty acid
VS	Volatile Solid
VSS	Volatile Suspended Solid
Y_{max}	Yield

References

Abbasi, T., Tauseef, S.M. and Abbasi, S.A., 2012. A brief history of anaerobic digestion and "biogas". In Biogas energy (pp. 11-23). Springer, New York, NY.

Abendroth, C., Vilanova, C., Günther, T., Luschnig, O. and Porcar, M., 2015. Eubacteria and archaea communities in seven mesophile anaerobic digester plants in Germany. Biotechnology for biofuels, 8(1), p.87.

Abowei, M.F.N., Ayotamuno, M.J. and Eze, C.L., 2009. Comparative evaluation of batch and continuous anaerobic digesters in biogas production from municipal solid waste using mathematical models. Agricultural Engineering International: CIGR Journal.

Adney, W.S., Rivard, C.J., Shiang, M. and Himmel, M.E., 1991. Anaerobic digestion of lignocellulosic biomass and wastes. Applied biochemistry and biotechnology, 30(2), pp.165-183.

Agler, M.T., Wrenn, B.A., Zinder, S.H. and Angenent, L.T., 2011. Waste to bioproduct conversion with undefined mixed cultures: the carboxylate platform. Trends in biotechnology, 29(2), pp.70-78.

Agne, M., Estelmann, S., Seelmann, C.S., Kung, J., Wilkens, D., Koch, H.G., van der Does, C., Albers, S.V., von Ballmoos, C., Simon, J. and Boll, M., 2021. The missing enzymatic link in syntrophic methane formation from fatty acids. Proceedings of the National Academy of Sciences, 118(40).

Agnihotri, G. and Liu, H.W., 2003. Enoyl-CoA hydratase: reaction, mechanism, and inhibition. Bioorganic & medicinal chemistry, 11(1), pp.9-20.

Ahring, B.K., Sandberg, M. and Angelidaki, I.J.A.M., 1995. Volatile fatty acids as indicators of process imbalance in anaerobic digestors. Applied microbiology and biotechnology, 43(3), pp.559-565.

Al Seadi, T., Drosg, B., Fuchs, W., Rutz, D. and Janssen, R., 2013. Biogas digestate quality and utilization. In The biogas handbook (pp. 267-301). Woodhead Publishing.

Allison, M.J., 1978. Production of branched-chain volatile fatty acids by certain anaerobic bacteria. Applied and environmental microbiology, 35(5), pp.872-877.

Alneberg, J., Bjarnason, B.S., De Bruijn, I., Schirmer, M., Quick, J., Ijaz, U.Z., Lahti, L., Loman, N.J., Andersson, A.F. and Quince, C., 2014. Binning metagenomic contigs by coverage and composition. Nature methods, 11(11), pp.1144-1146.

Angelidaki, I. and Ahring, B.K., 1992. Effects of free long-chain fatty acids on thermophilic anaerobic digestion. Applied microbiology and biotechnology, 37(6), pp.808-812.

Angelidaki, I., Karakashev, D., Batstone, D.J., Plugge, C.M. and Stams, A.J., 2011. Biomethanation and its potential. In Methods in enzymology (Vol. 494, pp. 327-351). Academic Press.

Angelonidi, E. and Smith, S.R., 2015. A comparison of wet and dry anaerobic digestion processes for the treatment of municipal solid waste and food waste. Water and environment journal, 29(4), pp.549-557.

Angenent, L.T., Karim, K., Al-Dahhan, M.H., Wrenn, B.A. and Domíguez-Espinosa, R., 2004. Production of bioenergy and biochemicals from industrial and agricultural wastewater. TRENDS in Biotechnology, 22(9), pp.477-485.

Arnosti, C., 2011. Microbial extracellular enzymes and the marine carbon cycle. Annual review of marine science, 3, pp.401-425.

Arnosti, C., Reintjes, G. and Amann, R., 2018. A mechanistic microbial underpinnin for the size-reactivity continuum of dissolved organic carbon degradation. Marine Chemistry, 206, pp.93-99.

Bagert, J.D., Xie, Y.J., Sweredoski, M.J., Qi, Y., Hess, S., Schuman, E.M. and Tirrell, D.A., 2014. Quantitative, time-resolved proteomic analysis by combining bioorthogonal noncanonical amino acid tagging and pulsed stable isotope labeling by amino acids in cell culture. Molecular & Cellular Proteomics, 13(5), pp.1352-1358.

Barker, H.A., 1936. Studies upon the methane-producing bacteria. Archiv für Mikrobiologie, 7(1-5), pp.420-438.

Barlow, J.T., Bogatyrev, S.R. and Ismagilov, R.F., 2020. A quantitative sequencing framework for absolute abundance measurements of mucosal and lumenal microbial communities. Nature communications, 11(1), pp.1-13.

Bayer, E.A., Shimon, L.J., Shoham, Y. and Lamed, R., 1998. Cellulosomes—structure and ultrastructure. Journal of structural biology, 124(2-3), pp.221-234.

Bernard, L., Mougel, C., Maron, P.A., Nowak, V., Lévêque, J., Henault, C., Haichar, F.E.Z., Berge, O., Marol, C., Balesdent, J. and Gibiat, F., 2007. Dynamics and identification of soil microbial populations actively assimilating carbon from 13C-labelled wheat residue as estimated by DNA- and RNA-SIP techniques. Environmental Microbiology, 9(3), pp.752-764.

Berry, D., Mader, E., Lee, T.K., Woebken, D., Wang, Y., Zhu, D., Palatinszky, M., Schintlmeister, A., Schmid, M.C., Hanson, B.T. and Shterzer, N., 2015. Tracking heavy water (D2O) incorporation for identifying and sorting active microbial cells. Proceedings of the National Academy of Sciences, 112(2), pp.E194-E203.

Blackburn, T.H., 1968. Protease production by Bacteroides amylophilus strain H 18. Microbiology, 53(1), pp.27-36.

Blum, D.L., Kataeva, I.A., Li, X.L. and Ljungdahl, L.G., 2000. Feruloyl esterase activity of the Clostridium thermocellum cellulosome can be attributed to

previously unknown domains of XynY and XynZ. Journal of bacteriology, 182(5), pp.1346-1351.

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. and Bai, Y., 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature biotechnology, 37(8), pp.852-857.

Bond, T. and Templeton, M.R., 2011. History and future of domestic biogas plants in the developing world. Energy for Sustainable development, 15(4), pp.347-354.

Boone, D.R. and Bryant, M.P., 1980. Propionate-degrading bacterium, Syntrophobacter wolinii sp. nov. gen. nov., from methanogenic ecosystems. Applied and environmental microbiology, 40(3), pp.626-632.

Boone, D.R., 1985. Fermentation reactions of anaerobic digestion. In PN Cheremisinoff and RP Oullette (edn), Biotechnology: applications and research. Technomic Publishing Co., Lancaster, PA. pp. 41-51.

Borneman, J., 1999. Culture-independent identification of microorganisms that respond to specified stimuli. Applied and environmental microbiology, 65(8), pp.3398-3400.

Bradford, L.M., Vestergaard, G., Táncsics, A., Zhu, B., Schloter, M. and Lueders, T., 2018. Transcriptome-stable isotope probing provides targeted functional and taxonomic insights into microaerobic pollutant-degrading aquifer microbiota. Frontiers in microbiology, 9, p.2696.

British Petroleum Company. BP Statistical Review of World Energy. London: British Petroleum Co, 2020.

Brondz, I., 2002. Development of fatty acid analysis by high-performance liquid chromatography, gas chromatography, and related techniques. Analytica Chimica Acta, 465(1-2), pp.1-37.

Brondz, I., 2016. LIPIDS, Fatty Acids, Encyclopedia of Analytical Science. In Chemistry, molecular sciences and chemical engineering. Elsevier Inc.

Brooks, J.P., 2016. Challenges for case-control studies with microbiome data. Annals of epidemiology, 26(5), pp.336-341.

Brown, T.A., 2020. Gene cloning and DNA analysis: an introduction. John Wiley & Sons.

Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K. and Earl, A.M., 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PloS one, 9(11), p.e112963.

Cai, M., Wilkins, D., Chen, J., Ng, S.K., Lu, H., Jia, Y. and Lee, P.K., 2016. Metagenomic reconstruction of key anaerobic digestion pathways in municipal sludge and industrial wastewater biogas-producing systems. Frontiers in Microbiology, 7, p.778.

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), pp.2639-2643.

Callahan, B.J., Sankaran, K., Fukuyama, J.A., McMurdie, P.J. and Holmes, S.P., 2016. Bioconductor workflow for microbiome data analysis: from raw reads to community analyses. F1000Research, 5.

Campanaro, S., Treu, L., Kougias, P.G., De Francisci, D., Valle, G. and Angelidaki, I., 2016. Metagenomic analysis and functional characterization of the biogas microbiome using high throughput shotgun sequencing and a novel binning strategy. Biotechnology for biofuels, 9(1), p.26.

Campanaro, S., Treu, L., Kougias, P.G., Luo, G. and Angelidaki, I., 2018. Metagenomic binning reveals the functional roles of core abundant microorganisms in twelve full-scale biogas plants. Water research, 140, pp.123-134.

Cangelosi, G.A. and Meschke, J.S., 2014. Dead or alive: molecular assessment of microbial viability. Applied and environmental microbiology, 80(19), pp.5884-5891.

Carballa, M., Regueiro, L. and Lema, J.M., 2015. Microbial management of anaerobic digestion: exploiting the microbiome-functionality nexus. Current opinion in biotechnology, 33, pp.103-111.

Cavaleiro, A.J., Sousa, D.Z., Alves, M.M., 2010. Methane production from oleate: Assessing the bioaugmentation potential of Syntrophomonas zehnderi. Water Research. 44, 4940-4947.

Chemerys, A., Pelletier, E., Cruaud, C., Martin, F., Violet, F. and Jouanneau, Y., 2014. Characterization of novel polycyclic aromatic hydrocarbon dioxygenases from the bacterial metagenomic DNA of a contaminated soil. Applied and environmental microbiology, 80(21), pp.6591-6600.

Chen, L., Zhao, B., Li, X., Cheng, Z. and Xia, Y., 2021. Isolating and characterizing translationally active fraction of anammox microbiota using bioorthogonal non-canonical amino acid tagging. Chemical Engineering Journal, 418, p.129411.

Chen, S., Liu, X., Dong, X., 2005. Syntrophobacter sulfatireducens sp. nov., a novel syntrophic, propionate-oxidizing bacterium isolated from UASB reactors. International Journal of Systematic and Evolutionary Microbiology. 55, 1319-1324. DOI 10.1099/ijs.0.63565-0

Chen, T.H. and Hashimoto, A.G., 1996. Effects of pH and substrate: inoculum ratio on batch methane fermentation. Bioresource technology, 56(2-3), pp.179-186.

Chen, Y., Dumont, M.G., Neufeld, J.D., Bodrossy, L., Stralis-Pavese, N., McNamara, N.P., Ostle, N., Briones, M.J. and Murrell, J.C., 2008. Revealing the uncultivated majority: combining DNA stable-isotope probing, multiple displacement amplification and metagenomic analyses of uncultivated

Methylocystis in acidic peatlands. Environmental Microbiology, 10(10), pp.2609-2622.

Chin, C.S., Peluso, P., Sedlazeck, F.J., Nattestad, M., Concepcion, G.T., Clum, A., Dunn, C., O'Malley, R., Figueroa-Balderas, R., Morales-Cruz, A. and Cramer, G.R., 2016. Phased diploid genome assembly with single-molecule real-time sequencing. Nature methods, 13(12), pp.1050-1054.

Chun-Te Lin, J., Liu, Y.S. and Wang, W.K., 2020. A full-scale study of high-rate anaerobic bioreactors for whiskey distillery wastewater treatment with size fractionation and metagenomic analysis of granular sludge. Bioresource Technology, 306, p.123032.

Cioabla, A.E., Ionel, I., Dumitrel, G.A. and Popescu, F., 2012. Comparative study on factors affecting anaerobic digestion of agricultural vegetal residues. Biotechnology for biofuels, 5(1), pp.1-9.

Cirne, D.G., Delgado, O.D., Marichamy, S. and Mattiasson, B., 2006. Clostridium lundense sp. nov., a novel anaerobic lipolytic bacterium isolated from bovine rumen. International journal of systematic and evolutionary microbiology, 56(3), pp.625-628.

Clark, D.P. and Cronan, J.E., 2005. Two-carbon compounds and fatty acids as carbon sources. EcoSal Plus, 1(2).

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), pp.486-496.

Couradeau, E., Sasse, J., Goudeau, D., Nath, N., Hazen, T.C., Bowen, B.P., Chakraborty, R., Malmstrom, R.R. Northen, T.R., 2019. Probing the active fraction of soil microbiomes using BONCAT-FACS. Nature Communications. 10, 2770.

Cuskin, F., Lowe, E.C., Temple, M.J., Zhu, Y., Cameron, E.A., Pudlo, N.A., Porter, N.T., Urs, K., Thompson, A.J., Cartmell, A. and Rogowski, A., 2015. Human gut Bacteroidetes can utilize yeast mannan through a selfish mechanism. Nature, 517(7533), pp.165-169.

Dahiya, A. ed., 2014. Bioenergy: Biomass to biofuels. Academic Press.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M. and Li, H., 2021. Twelve years of SAMtools and BCFtools. Gigascience, 10(2), p.giab008.

David, M.M., Cecillon, S., Warne, B.M., Prestat, E., Jansson, J.K. and Vogel, T.M., 2015. Microbial ecology of chlorinated solvent biodegradation. Environmental microbiology, 17(12), pp.4835-4850.

de Bok, F.A., Stams, A.J., Dijkema, C. and Boone, D.R., 2001. Pathway of propionate oxidation by a syntrophic culture of Smithella propionica and Methanospirillum hungatei. Applied and environmental microbiology, 67(4), pp.1800-1804.

de Bok, F.A.M., Harmsen, H.J.M., Plugge, C.M., de Vries, M.C., Akkermans, A.D.L., de Vos, W.M., and Stams, A.J.M., 2005. The first true obligately syntrophic propionateoxidizing bacterium, Pelotomaculum schinkii sp. nov., co-cultured with Methanospirillum hungatei, and emended description of the genus Pelotomaculum. International Journal of Systematic and Evolutionary Microbiology 55, 1697-1703

De Francisci, D., Kougias, P.G., Treu, L., Campanaro, S. and Angelidaki, I., 2015. Microbial diversity and dynamicity of biogas reactors due to radical changes of feedstock composition. Bioresource technology, 176, pp.56-64.

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), pp.2797-2809.

De Vrieze, J., Saunders, A.M., He, Y., Fang, J., Nielsen, P.H., Verstraete, W. and Boon, N., 2015. Ammonia and temperature determine potential clustering in the anaerobic digestion microbiome. Water research, 75, pp.312-323.

Desvaux, M., 2005. The cellulosome of Clostridium cellulolyticum. Enzyme and Microbial Technology, 37(4), pp.373-385.

Díaz, A.I., Oulego, P., Collado, S., Laca, A., González, J.M. and Díaz, M., 2018. Impact of anaerobic digestion and centrifugation/decanting processes in bacterial communities fractions. Journal of bioscience and bioengineering, 126(6), pp.742-749.

Dieterich, D.C. et al., 2006. Selective identification of newly synthesized proteins in mammalian cells using biorthogonal noncanonical amino acid tagging (BONCAT). Proc. Natl. Acad. Sci. USA. 103:9482-9487

Dieterich, D.C., Lee, J.J., Link, A.J., Graumann, J., Tirrell, D.A. and Schuman, E.M., 2007. Labeling, detection and identification of newly synthesized proteomes with bioorthogonal non-canonical amino-acid tagging. Nature protocols, 2(3), pp.532-540.

DiRusso, C.C., Black, P.N., Weimar, J.D., 1999. Molecular inroads into the regulation and metabolism of fatty acids, lessons from bacteria. Prog Lipid Res. 38, 129–197.

Doi, R.H., 2008. Cellulases of mesophilic microorganisms: cellulosome and noncellulosome producers. Annals of the New York Academy of Sciences, 1125(1), pp.267-279.

Dong, Y., Liu, Y., Chen, N., Zhong, Y., Liu, L. and Xie, Q., 2018. Clostridium beihaiense sp. nov., an anaerobic bacterium isolated from activated sludge. International journal of systematic and evolutionary microbiology, 68(9), pp.2789-2793.

Dong, W., Song, A., Yin, H., Liu, X., Li, J. and Fan, F., 2021. The decomposition of microbial necromass is divergent at the individual taxonomic level in soil. Frontiers in microbiology, 12, p.1674.

Dowle, E.J., Pochon, X., C. Banks, J., Shearer, K. and Wood, S.A., 2016. Targeted gene enrichment and high-throughput sequencing for environmental biomonitoring: A case study using freshwater macroinvertebrates. Molecular ecology resources, 16(5), pp.1240-1254.

Drake, H.L. and Daniel, S.L., 2004. Physiology of the thermophilic acetogen Moorella thermoacetica. Research in microbiology, 155(10), pp.869-883.

Drake, H.L., Gößner, A.S. and Daniel, S.L., 2008. Old acetogens, new light. Annals of the New York Academy of Sciences, 1125(1), pp.100-128.

Dubé, C.D. and Guiot, S.R., 2015. Direct interspecies electron transfer in anaerobic digestion: a review. Biogas science and technology, pp.101-115.

Dumont, M.G. and Murrell, J.C., 2005. Stable isotope probing—linking microbial identity to function. Nature Reviews Microbiology, 3(6), pp.499-504.

Dumont, M.G., Pommerenke, B., Casper, P. and Conrad, R., 2011. DNA-, rRNA- and mRNA-based stable isotope probing of aerobic methanotrophs in lake sediment. Environmental Microbiology, 13(5), pp.1153-1167.

Dworkin, J. and Shah, I.M., 2010. Exit from dormancy in microbial organisms. Nature reviews microbiology, 8(12), pp.890-896.

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics, 26(19), pp.2460-2461.

Edlund, A., Hårdeman, F., Jansson, J.K. and Sjöling, S., 2008. Active bacterial community structure along vertical redox gradients in Baltic Sea sediment. Environmental Microbiology, 10(8), pp.2051-2063.

Eichorst, S.A., Strasser, F., Woyke, T., Schintlmeister, A., Wagner, M. and Woebken, D., 2015. Advancements in the application of NanoSIMS and Raman microspectroscopy to investigate the activity of microbial cells in soils. FEMS microbiology ecology, 91(10), p.fiv106.

Emerson, J.B., Adams, R.I., Román, C.M.B., Brooks, B., Coil, D.A., Dahlhausen, K., Ganz, H.H., Hartmann, E.M., Hsu, T., Justice, N.B. and Paulino-Lima, I.G., 2017. Schrödinger's microbes: tools for distinguishing the living from the dead in microbial ecosystems. Microbiome, 5(1), pp.1-23.

Enzmann, F., Mayer, F., Rother, M. and Holtmann, D., 2018. Methanogens: biochemical background and biotechnological applications. AMB Express, 8(1), pp.1-22.

Eren, A.M., Maignien, L., Sul, W.J., Murphy, L.G., Grim, S.L., Morrison, H.G. and Sogin, M.L., 2013. Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. Methods in ecology and evolution, 4(12), pp.1111-1119.

Eren, A.M., Morrison, H.G., Lescault, P.J., Reveillaud, J., Vineis, J.H. and Sogin, M.L., 2015. Minimum entropy decomposition: unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. The ISME journal, 9(4), pp.968-979.

Eyice, Ö., Namura, M., Chen, Y., Mead, A., Samavedam, S. and Schäfer, H., 2015. SIP metagenomics identifies uncultivated Methylophilaceae as dimethylsulphide degrading bacteria in soil and lake sediment. The ISME journal, 9(11), pp.2336-2348.

Feitkenhauer, H., von Sachs, J. and Meyer, U., 2002. On-line titration of volatile fatty acids for the process control of anaerobic digestion plants. Water Research, 36(1), pp.212-218.

Fernández, A., Sanchez, A. and Font, X., 2005. Anaerobic co-digestion of a simulated organic fraction of municipal solid wastes and fats of animal and vegetable origin. Biochemical Engineering Journal, 26(1), pp.22-28.

Ferry, J.G., 2010. The chemical biology of methanogenesis. Planetary and Space Science, 58(14-15), pp.1775-1783.

Fierer, N., Leff, J.W., Adams, B.J., Nielsen, U.N., Bates, S.T., Lauber, C.L., Owens, S., Gilbert, J.A., Wall, D.H. and Caporaso, J.G., 2012. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. Proceedings of the National Academy of Sciences, 109(52), pp.21390-21395.

Flemming, H.C. and Wuertz, S., 2019. Bacteria and archaea on Earth and their abundance in biofilms. Nature Reviews Microbiology, 17(4), pp.247-260.

Fortunato, C.S. and Huber, J.A., 2016. Coupled RNA-SIP and metatranscriptomics of active chemolithoautotrophic communities at a deep-sea hydrothermal vent. The ISME journal, 10(8), pp.1925-1938.

Freguia S., Virdis B., Rabaey K., 2013. Biofuels. In: Rosenberg E., DeLong E.F., Lory S., Stackebrandt E., Thompson F. (eds) The Prokaryotes. Springer, Berlin, Heidelberg.

Fu, B., Conrad, R. and Blaser, M., 2018. Potential contribution of acetogenesis to anaerobic degradation in methanogenic rice field soils. Soil Biology and Biochemistry, 119, pp.1-10.

Galivan, J.H. and Allen, S.H.G., 1968. Methylmalonyl coenzyme A decarboxylase: its role in succinate decarboxylation by Micrococcus lactilyticus. Journal of Biological Chemistry, 243(6), pp.1253-1261.

Ganesh, R., Torrijos, M., Sousbie, P., Lugardon, A., Steyer, J.P. and Delgenes, J.P., 2014. Single-phase and two-phase anaerobic digestion of fruit and vegetable waste: comparison of start-up, reactor stability and process performance. Waste management, 34(5), pp.875-885.

Garcia, J.L., Patel, B.K. and Ollivier, B., 2000. Taxonomic, phylogenetic, and ecological diversity of methanogenic Archaea. Anaerobe, 6(4), pp.205-226.

Ghidotti, M., Fabbri, D., Torri, C. and Piccinini, S., 2018. Determination of volatile fatty acids in digestate by solvent extraction with dimethyl carbonate and gas chromatography-mass spectrometry. Analytica chimica acta, 1034, pp.92-101.

Gibson, S.A.W. and Macfarlane, G.T., 1988. Characterization of proteases formed by Bacteroides fragilis. Microbiology, 134(8), pp.2231-2240.

Gibson, B., Wilson, D.J., Feil, E. and Eyre-Walker, A., 2018. The distribution of bacterial doubling times in the wild. Proceedings of the Royal Society B, 285(1880), p.20180789.

Goberna, M., Insam, H. and Franke-Whittle, I., 2009. Effect of biowaste sludge maturation on the diversity of thermophilic bacteria and archaea in an anaerobic reactor. Applied and environmental microbiology, 75(8), pp.2566-2572.

Goldfarb, K.C., Karaoz, U., Hanson, C.A., Santee, C.A., Bradford, M.A., Treseder, K.K., Wallenstein, M.D. and Brodie, E.L., 2011. Differential growth responses of soil bacterial taxa to carbon substrates of varying chemical recalcitrance. Frontiers in microbiology, 2, p.94.

Gourdon, R. and Vermande, P., 1987. Effects of propionic acid concentration on anaerobic digestion of pig manure. Biomass, 13(1), pp.1-12.

Green, J.L., Bohannan, B.J. and Whitaker, R.J., 2008. Microbial biogeography: from taxonomy to traits. science, 320(5879), pp.1039-1043.

Griffin, R.J., 1994. 3 The Medicinal Chemistry of the Azido Group. Progress in medicinal chemistry, 31, pp.121-232.

Gujer, W. and Zehnder, A.J., 1983. Conversion processes in anaerobic digestion. Water science and technology, 15(8-9), pp.127-167.

Gülay, A., Fowler, S.J., Tatari, K., Thamdrup, B., Albrechtsen, H.J., Al-Soud, W.A., Sørensen, S.J. and Smets, B.F., 2019. DNA-and RNA-SIP reveal Nitrospira spp. as key drivers of nitrification in groundwater-fed biofilters. MBio, 10(6), pp.e01870-19.

Gundawar, K., Kumari, S., Sharma, S., Grover, V., Patil, P.B. and Korpole, S., 2019. Clostridium indicum sp. nov., a novel anaerobic bacterium isolated from the sludge of an industrial effluent. International journal of systematic and evolutionary microbiology, 69(3), pp.672-678.

Guo, J., Peng, Y., Ni, B.J., Han, X., Fan, L. and Yuan, Z., 2015. Dissecting microbial community structure and methane-producing pathways of a full-scale anaerobic reactor digesting activated sludge from wastewater treatment by metagenomic sequencing. Microbial cell factories, 14(1), p.33.

Gutierrez, T., Singleton, D.R., Berry, D., Yang, T., Aitken, M.D. and Teske, A., 2013. Hydrocarbon-degrading bacteria enriched by the Deepwater Horizon oil spill identified by cultivation and DNA-SIP. The ISME journal, 7(11), pp.2091-2104.

Hamasaki, K., Taniguchi, A., Tada, Y., Kaneko, R. and Miki, T., 2016. Active populations of rare microbes in oceanic environments as revealed by bromodeoxyuridine incorporation and 454 tag sequencing. Gene, 576(2), pp.650-656.

Hanreich, A., Heyer, R., Benndorf, D., Rapp, E., Pioch, M., Reichl, U. and Klocke, M., 2012. Metaproteome analysis to determine the metabolically active part of a thermophilic microbial community producing biogas from agricultural biomass. Canadian journal of microbiology, 58(7), pp.917-922.

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), pp.330-338.

Harmsen, H.J., Van Kuijk, B.L., Plugge, C.M., Akkermans, A.D., De Vos, W.M. and Stams, A.J., 1998. Syntrophobacter fumaroxidans sp. nov., a syntrophic propionate-degrading sulfate-reducing bacterium. International journal of systematic and evolutionary microbiology, 48(4), pp.1383-1387.

Harvey, K.L., Jarocki, V.M., Charles, I.G. and Djordjevic, S.P., 2019. The diverse functional roles of elongation factor Tu (EF-Tu) in microbial pathogenesis. Frontiers in microbiology, 10, p.2351.

Hassa, J., Maus, I., Off, S., Pühler, A., Scherer, P., Klocke, M. and Schlüter, A., 2018. Metagenome, metatranscriptome, and metaproteome approaches unraveled compositions and functional relationships of microbial communities residing in biogas plants. Applied microbiology and biotechnology, 102(12), pp.5045-5063.

Hatamoto, M., Imachi, H., Fukayo, S., Ohashi, A., Harada, H., 2007. Syntrophomonas palmitatica sp. nov., an anaerobic, syntrophic, long-chain fatty-acid-oxidizing bacterium isolated from methanogenic sludge. International Journal of Systematic and Evolutionary Microbiology. 57, 2137-2142.

Hatamoto, M., Kaneshige, M., Nakamura, A. and Yamaguchi, T., 2014. Bacteroides luti sp. nov., an anaerobic, cellulolytic and xylanolytic bacterium isolated from methanogenic sludge. International journal of systematic and evolutionary microbiology, 64(5), pp.1770-1774.

Hatzenpichler, R. and Orphan, V.J., 2015. Detection of protein-synthesizing microorganisms in the environment via bioorthogonal noncanonical amino acid tagging (BONCAT). In Hydrocarbon and Lipid Microbiology Protocols (pp. 145-157). Springer, Berlin, Heidelberg.

Hatzenpichler, R., Connon, S.A., Goudeau, D., Malmstrom, R.R., Woyke, T., and Orphan, V.J., 2016. Visualizing in situ translational activity for identifying and sorting slow-growing archaeal-bacterial consortia. PNAS 113(28), 4069-4078

Hatzenpichler, R., Krukenberg, V., Spietz, R.L. and Jay, Z.J., 2020. Next-generation physiology approaches to study microbiome function at single cell level. Nature Reviews Microbiology, pp.1-16.

Hatzenpichler, R., Scheller, S., Tavormina, P.L., Babin, B.M., Tirrell, D.A., and Orphan, V.J., 2014. In situ visualization of newly synthesized proteins in environmental microbes using amino acid tagging and click chemistry. Envirionmental Microbiology 16, 2568-2590

Hausmann, B., Pelikan, C., Rattei, T., Loy, A. and Pester, M., 2019. Long-term transcriptional activity at zero growth of a cosmopolitan rare biosphere member. MBio, 10(1), pp.e02189-18.

Hazlewood, G.P. and Edwards, R., 1981. Proteolytic activities of a rumen bacterium, Bacteroides ruminicola R8/4. Microbiology, 125(1), pp.11-15.

Heeg, K., Pohl, M., Sontag, M., Mumme, J., Klocke, M. and Nettmann, E., 2014. Microbial communities involved in biogas production from wheat straw as the sole substrate within a two-phase solid-state anaerobic digestion. Systematic and applied microbiology, 37(8), pp.590-600.

Henze, M. and Harremoës, P., 1983. Anaerobic treatment of wastewater in fixed film reactors–a literature review. Water science and technology, 15(8-9), pp.1-101.

Hernández-Eugenio, G., Fardeau, M.L., Cayol, J.L., Patel, B.K., Thomas, P., Macarie, H., Garcia, J.L. and Ollivier, B., 2002. Clostridium thiosulfatireducens sp. nov., a proteolytic, thiosulfate-and sulfur-reducing bacterium isolated from an upflow anaerobic sludge blanket (UASB) reactor. International journal of systematic and evolutionary microbiology, 52(5), pp.1461-1468.

Heukelekian, H. and Mueller, P., 1958. Transformation of some lipids in anaerobic sludge digestion. Sewage and industrial wastes, 30(9), pp.1108-1120.

Hill, C.G. and Root, T.W., 1977. An introduction to chemical engineering kinetics & reactor design (pp. 245-316). New York: Wiley.

Hill, D.T. and Holmberg, R.D., 1988. Long chain volatile fatty acid relationships in anaerobic digestion of swine waste. Biological wastes, 23(3), pp.195-214.

Holmes, D.E., Shrestha, P.M., Walker, D.J., Dang, Y., Nevin, K.P., Woodard, T.L. and Lovley, D.R., 2017. Metatranscriptomic evidence for direct interspecies electron transfer between Geobacter and Methanothrix species in methanogenic rice paddy soils. Applied and environmental microbiology, 83(9), pp.e00223-17.

Huang, W.E., Stoecker, K., Griffiths, R., Newbold, L., Daims, H., Whiteley, A.S. and Wagner, M., 2007. Raman-FISH: combining stable-isotope Raman spectroscopy and fluorescence in situ hybridization for the single cell analysis of identity and function. Environmental microbiology, 9(8), pp.1878-1889.

Huisgen, R., 1963. 1, 3-dipolar cycloadditions. Past and future. Angewandte Chemie International Edition in English, 2(10), pp.565-598.

Huser, B.A., Wuhrmann, K. and Zehnder, A.J., 1982. Methanothrix soehngenii gen. nov. sp. nov., a new acetotrophic non-hydrogen-oxidizing methane bacterium. Archives of Microbiology, 132(1), pp.1-9.

Hyatt, D., Chen, G.L., LoCascio, P.F., Land, M.L., Larimer, F.W. and Hauser, L.J., 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC bioinformatics, 11(1), pp.1-11.

IEA. 2020. World Energy Statistics 2019. IEA, Paris. https://www.iea.org/reports/world-energy-statistics-2019

Iino, T., Tamaki, H., Tamazawa, S., Ueno, Y., Ohkuma, M., Suzuki, K.I., Igarashi, Y. and Haruta, S., 2013. Candidatus Methanogranum caenicola: a novel methanogen from the anaerobic digested sludge, and proposal of Methanomassiliicoccaceae fam. nov. and Methanomassiliicoccales ord. nov., for a methanogenic lineage of the class Thermoplasmata. Microbes and environments, p.ME12189.

Imachi, H., Sakai, S., Ohashi, A., Harada, H., Hanada, S., Kamagata, Y. and Sekiguchi, Y., 2007. Pelotomaculum propionicicum sp. nov., an anaerobic, mesophilic, obligatelysyntrophic, propionate-oxidizing bacterium. Int J Syst Evol Microbiol 57, 1487-1492.

Imachi, H., Sekiguchi, Y., Kamagata, Y., Hanada, S., Ohashi, A., Harada, H., 2002. Pelotomaculum thermopropionicum gen. nov., sp. nov., an anaerobic, thermophilic, syntrophic propionate-oxidizing bacterium. International Journal of Systematic and Evolutionary Microbiology. 52, 1729-1735.

Imachi, H., Sakai, S., Sekiguchi, Y., Hanada, S., Kamagata, Y., Ohashi, A. and Harada, H., 2008. Methanolinea tarda gen. nov., sp. nov., a methane-producing archaeon isolated from a methanogenic digester sludge. International Journal of Systematic and Evolutionary Microbiology, 58(1), pp.294-301.

Iwai, S., Weinmaier, T., Schmidt, B.L., Albertson, D.G., Poloso, N.J., Dabbagh, K. and DeSantis, T.Z., 2016. Piphillin: improved prediction of metagenomic content by direct inference from human microbiomes. PloS one, 11(11), p.e0166104.

Jackson, B.E., Bhupathiraju, V.K., Tanner, R.S., Woese, C.R. and McInerney, M.J., 1999. Syntrophus aciditrophicus sp. nov., a new anaerobic bacterium that degrades fatty acids and benzoate in syntrophic association with hydrogen-using microorganisms. Arch Microbiol 171, 107-114.

Jain, M.K. and Zeikus, J.G., 1988. Taxonomic distinction of two new protein specific, hydrolytic anaerobes: Isolation and characterization of Clostridium proteolyticum sp. nov. and Clostridium collagenovorans sp. nov. Systematic and applied microbiology, 10(2), pp.134-141.

James, A.G., Austin, C.J., Cox, D.S., Taylor, D. and Calvert, R., 2013. Microbiological and biochemical origins of human axillary odour. FEMS microbiology ecology, 83(3), pp.527-540.

Janda, J.M. and Abbott, S.L., 2007. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. Journal of clinical microbiology, 45(9), pp.2761-2764.

Jeganathan, J., Nakhla, G. and Bassi, A., 2006. Long-term performance of high-rate anaerobic reactors for the treatment of oily wastewater. Environmental science & technology, 40(20), pp.6466-6472.

Jimenez-Diaz, L., Caballero, A. and Segura, A., 2017. Pathways for the degradation of fatty acids in bacteria. Aerobic utilization of hydrocarbons, oils and lipids, 10, pp.978-3.

Johnson, J.S., Spakowicz, D.J., Hong, B.Y., Petersen, L.M., Demkowicz, P., Chen, L., Leopold, S.R., Hanson, B.M., Agresta, H.O., Gerstein, M. and Sodergren, E., 2019. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nature communications, 10(1), pp.1-11.

Jun, S.R., Robeson, M.S., Hauser, L.J., Schadt, C.W. and Gorin, A.A., 2015. PanFP: pangenome-based functional profiles for microbial communities. BMC research notes, 8(1), pp.1-7.

Kalesh, K. and Denny, P.W., 2019. A BONCAT-iTRAQ method enables temporally resolved quantitative profiling of newly synthesised proteins in Leishmania mexicana parasites during starvation. PLoS neglected tropical diseases, 13(12), p.e0007651.

Käll, L., Storey, J.D. and Noble, W.S., 2008. Non-parametric estimation of posterior error probabilities associated with peptides identified by tandem mass spectrometry. Bioinformatics, 24(16), pp.i42-i48.

Kalyuzhnaya, M.G., Lapidus, A., Ivanova, N., Copeland, A.C., McHardy, A.C., Szeto, E., Salamov, A., Grigoriev, I.V., Suciu, D., Levine, S.R. and Markowitz, V.M., 2008. High-resolution metagenomics targets specific functional types in complex microbial communities. Nature biotechnology, 26(9), pp.1029-1034.

Kamp, F. and Hamilton, J.A., 2006. How fatty acids of different chain length enter and leave cells by free diffusion. Prostaglandins, leukotrienes and essential fatty acids, 75(3), pp.149-159.

Kampmann, K., Ratering, S., Kramer, I., Schmidt, M., Zerr, W. and Schnell, S., 2012. Unexpected stability of Bacteroidetes and Firmicutes communities in laboratory biogas reactors fed with different defined substrates. Appl. Environ. Microbiol., 78(7), pp.2106-2119.

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. and Tanabe, M., 2016. KEGG as a reference resource for gene and protein annotation. Nucleic acids research, 44(D1), pp.D457-D462.

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, p.e7359.

Karakashev, D., Batstone, D.J. and Angelidaki, I., 2005. Influence of environmental conditions on methanogenic compositions in anaerobic biogas reactors. Appl. Environ. Microbiol., 71(1), pp.331-338.

Kataeva, I.A., Seidel, R.D., Shah, A., West, L.T., Li, X.L. and Ljungdahl, L.G., 2002. The fibronectin type 3-like repeat from the Clostridium thermocellum cellobiohydrolase CbhA promotes hydrolysis of cellulose by modifying its surface. Appl. Environ. Microbiol., 68(9), pp.4292-4300.

Kendall, M.M., Liu, Y., Boone, D.R., 2006. Butyrate- and propionate-degrading syntrophs from permanently cold marine sediments in Skan Bay, Alaska, and description of Algorimarina butyrica gen. nov., sp. nov. FEMS Microbiol. 262, 107-114.

Kern, T., Theiss, J., Röske, K. and Rother, M., 2016. Assessment of hydrogen metabolism in commercial anaerobic digesters. Applied microbiology and biotechnology, 100(10), pp.4699-4710.

Khotsena, C. and Potivichayanon, S., 2020, October. Determination of Appropriate Conditions for Volatile Fatty Acids from Rubber Industrial Wastewater by GC-FID: Headspace Technique. In 2020 International Conference and Utility Exhibition on Energy, Environment and Climate Change (ICUE) (pp. 1-7). IEEE.

Kiick, K.L., Saxon, E., Tirrell, D.A. and Bertozzi, C.R., 2002. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. Proceedings of the National academy of Sciences, 99(1), pp.19-24.

Kiick, K.L. and Tirrell, D.A., 2000. Protein engineering by in vivo incorporation of non-natural amino acids: Control of incorporation of methionine analogues by methionyl-tRNA synthetase. Tetrahedron, 56(48), pp.9487-9493.

Kiick, K.L., Weberskirch, R. and Tirrell, D.A., 2001. Identification of an expanded set of translationally active methionine analogues in Escherichia coli. Febs Letters, 502(1-2), pp.25-30.

Kim, H., Kim, S. and Jung, S., 2020. Instruction of microbiome taxonomic profiling based on 16S rRNA sequencing. Journal of Microbiology, 58(3), pp.193-205.

Kim, M.S., Kim, D.H. and Yun, Y.M., 2017. Effect of operation temperature on anaerobic digestion of food waste: Performance and microbial analysis. Fuel, 209, pp.598-605.

Kirchman, D.L., 2018. Processes in Microbial Ecology. Second Edition. Oxford University Press.

Kirkegaard, R.H., McIlroy, S.J., Kristensen, J.M., Nierychlo, M., Karst, S.M., Dueholm, M.S., Albertsen, M. and Nielsen, P.H., 2017. The impact of immigration on microbial community composition in full-scale anaerobic digesters. Scientific reports, 7(1), pp.1-11.

Klocke, M., Mähnert, P., Mundt, K., Souidi, K. and Linke, B., 2007. Microbial community analysis of a biogas-producing completely stirred tank reactor fed continuously with fodder beet silage as mono-substrate. Systematic and applied microbiology, 30(2), pp.139-151.

Koch, C., Müller, S., Harms, H. and Harnisch, F., 2014. Microbiomes in bioenergy production: from analysis to management. Current opinion in biotechnology, 27, pp.65-72.

Kolmogorov, M., Bickhart, D.M., Behsaz, B., Gurevich, A., Rayko, M., Shin, S.B., Kuhn, K., Yuan, J., Polevikov, E., Smith, T.P. and Pevzner, P.A., 2020. metaFlye: scalable long-read metagenome assembly using repeat graphs. Nature Methods, 17(11), pp.1103-1110.

Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H. and Phillippy, A.M., 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome research, 27(5), pp.722-736.

Koropatkin, N.M., Cameron, E.A. and Martens, E.C., 2012. How glycan metabolism shapes the human gut microbiota. Nature Reviews Microbiology, 10(5), pp.323-335.

Koster, I.W. and Cramer, A., 1987. Inhibition of methanogenesis from acetate in granular sludge by long-chain fatty acids. Applied and environmental microbiology, 53(2), pp.403-409.

Kouzuma, A., Kato, S. and Watanabe, K., 2015. Microbial interspecies interactions: recent findings in syntrophic consortia. Frontiers in microbiology, 6, p.477.

Krehenwinkel, H., Wolf, M., Lim, J.Y., Rominger, A.J., Simison, W.B. and Gillespie, R.G., 2017. Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. Scientific reports, 7(1), pp.1-12.

Kurade, M.B., Saha, S., Salama, E.S., Patil, S.M., Govindwar, S.P. and Jeon, B.H., 2019. Acetoclastic methanogenesis led by Methanosarcina in anaerobic co-digestion of fats, oil and grease for enhanced production of methane. Bioresource technology, 272, pp.351-359.

Kurokawa, J., Hemjinda, E., Arai, T., Karita, S., Kimura, T., Sakka, K. and Ohmiya, K., 2001. Sequence of the Clostridium thermocellum mannanase gene man26B and characterization of the translated product. Bioscience, biotechnology, and biochemistry, 65(3), pp.548-554.

Lahav, O. and Morgan, B.E., 2004. Titration methodologies for monitoring of anaerobic digestion in developing countries—a review. Journal of Chemical Technology & Biotechnology: International Research in Process, Environmental & Clean Technology, 79(12), pp.1331-1341.

Lalman, J.A. and Bagley, D.M., 2000. Anaerobic degradation and inhibitory effects of linoleic acid. Water research, 34(17), pp.4220-4228.

Langer, S.G., Ahmed, S., Einfalt, D., Bengelsdorf, F.R. and Kazda, M., 2015. Functionally redundant but dissimilar microbial communities within biogas reactors treating maize silage in co-fermentation with sugar beet silage. Microbial biotechnology, 8(5), pp.828-836.

Langille, M.G., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., Clemente, J.C., Burkepile, D.E., Thurber, R.L.V., Knight, R. and Beiko, R.G., 2013. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nature biotechnology, 31(9), pp.814-821.

Laserson, J., Jojic, V. and Koller, D., 2011. Genovo: de novo assembly for metagenomes. Journal of Computational Biology, 18(3), pp.429-443.

Leach, A.L., Chong, J.P. and Redeker, K.R., 2012. SSuMMo: rapid analysis, comparison and visualization of microbial communities. Bioinformatics, 28(5), pp.679-686.

Lee, S.F., Forsberg, C.W. and Gibbins, L., 1985. Cellulolytic activity of Clostridium acetobutylicum. Applied and environmental microbiology, 50(2), pp.220-228.

Lee, W.H. and Riemann, H., 1970. The genetic relatedness of proteolytic Clostridium botulinum strains. Microbiology, 64(1), pp.85-90.

Lee, N., Nielsen, P.H., Andreasen, K.H., Juretschko, S., Nielsen, J.L., Schleifer, K.H. and Wagner, M., 1999. Combination of fluorescent in situ hybridization and microautoradiography—a new tool for structure-function analyses in microbial ecology. Applied and environmental microbiology, 65(3), pp.1289-1297.

Leizeaga, A., Estrany, M., Forn, I and Sebastian, M., 2017. Using click-chemistry for visualizing in situ changes of translational activity in planktonic marine bacteria. Front. Microbiol. 8: 2360.

Lemaire, O.N., Jespersen, M. and Wagner, T., 2020. CO2-fixation strategies in energy extremophiles: What can we learn from acetogens?. Frontiers in Microbiology, 11.

Li, A., Chu, Y.N., Wang, X., Ren, L., Yu, J., Liu, X., Yan, J., Zhang, L., Wu, S. and Li, S., 2013. A pyrosequencing-based metagenomic study of methane-producing microbial community in solid-state biogas reactor. Biotechnology for biofuels, 6(1), pp.1-17.

Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997.

Li, D., Liu, C.M., Luo, R., Sadakane, K. and Lam, T.W., 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics, 31(10), pp.1674-1676.

Li, H., 2016. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. Bioinformatics, 32(14), pp.2103-2110.

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.

Lin, C.Y. and Hu, Y.Y., 1993. Mesophilic degradation of butyric acid in anaerobic digestion. Journal of Chemical Technology & Biotechnology, 56(2), pp.191-194.

Link, A.J. and Tirrell, D.A., 2003. Cell surface labeling of escherichia c oli via copper (i)-catalyzed [3+ 2] cycloaddition. Journal of the American Chemical Society, 125(37), pp.11164-11165.

Liu, F.H., Wang, S.B., Zhang, J.S., Zhang, J., Yan, X., Zhou, H.K., Zhao, G.P. and Zhou, Z.H., 2009a. The structure of the bacterial and archaeal community in a biogas digester as revealed by denaturing gradient gel electrophoresis and 16S rDNA sequencing analysis. Journal of Applied Microbiology, 106(3), pp.952-966.

Liu, G., Zhang, R., El-Mashad, H.M. and Dong, R., 2009b. Effect of feed to inoculum ratios on biogas yields of food and green wastes. Bioresource technology, 100(21), pp.5103-5108.

Liu, X., Gao, X., Wang, W., Zheng, L., Zhou, Y. and Sun, Y., 2012. Pilot-scale anaerobic co-digestion of municipal biomass waste: Focusing on biogas production and GHG reduction. Renewable energy, 44, pp.463-468.

Liu, Y. and Whitman, W.B., 2008. Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. Annals of the New York Academy of Sciences, 1125(1), pp.171-189.

Liu, Y., Balkwill, D.L., Aldrich, H.C., Drake, G.R. and Boone, D.R., 1999. Characterization of the anaerobic propionate-degrading syntrophs Smithella propionica gen. nov., sp. nov. and Syntrophobacter wolinii. Int J Syst Bacteriol 49, 545-556.

Lloret, E., Pastor, L., Pradas, P. and Pascual, J.A., 2013. Semi full-scale thermophilic anaerobic digestion (TAnD) for advanced treatment of sewage sludge: stabilization process and pathogen reduction. Chemical Engineering Journal, 232, pp.42-50.

Loehr, R.C. and Roth, J.C., 1968. Aerobic degradation of long-chain fatty acid salts. Journal (Water Pollution Control Federation), pp.R385-R403.

Loman, N.J., Quick, J. and Simpson, J.T., 2015. A complete bacterial genome assembled de novo using only nanopore sequencing data. Nature methods, 12(8), pp.733-735.

Lopes, W.S., Leite, V.D. and Prasad, S., 2004. Influence of inoculum on performance of anaerobic reactors for treating municipal solid waste. Bioresource technology, 94(3), pp.261-266.

Lorowitz, W.H., Zhao, H., Bryant, M.P., 1989. Syntrophomonas wolfei subsp. saponavida subsp. nov. , a LongChain -Fatty- Acid-Degrading, Anaerobic, S yntrophic Bacterium; Syntrophomonas wolfei subsp. wolfei subsp. nov. ; and

Emended Descriptions of the Genus and Species. International Journal of Systematic Bacteriology. 39(2), 122-126.

Louca, S., Polz, M.F., Mazel, F., Albright, M.B., Huber, J.A., O'Connor, M.I., Ackermann, M., Hahn, A.S., Srivastava, D.S., Crowe, S.A. and Doebeli, M., 2018. Function and functional redundancy in microbial systems. Nature ecology & evolution, 2(6), pp.936-943.

Lukehurst, C.T., Frost, P. and Al Seadi, T., 2010. Utilisation of digestate from biogas plants as biofertiliser. IEA bioenergy, 2010, pp.1-36.

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y. and Tang, J., 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascience, 1(1), pp.2047-217X.

Luton, P.E., Wayne, J.M., Sharp, R.J. and Riley, P.W., 2002. The mcrA gene as an alternative to 16S rRNA in the phylogenetic analysis of methanogen populations in landfillbbThe GenBank accession numbers for the mcrA sequences reported in this paper are AF414034–AF414051 (see Fig. 2) and AF414007–AF414033 (environmental isolates in Fig. 3). Microbiology, 148(11), pp.3521-3530.

Lv, Z., Leite, A.F., Harms, H., Glaser, K., Liebetrau, J., Kleinsteuber, S. and Nikolausz, M., 2019. Microbial community shifts in biogas reactors upon complete or partial ammonia inhibition. Applied microbiology and biotechnology, 103(1), pp.519-533.

Ma, J., Frear, C., Wang, Z.W., Yu, L., Zhao, Q., Li, X. and Chen, S., 2013. A simple methodology for rate-limiting step determination for anaerobic digestion of complex substrates and effect of microbial community ratio. Bioresource technology, 134, pp.391-395.

Madigan, M.T., Bender, K.S., Buckley, D.H., Sattley, W.M. and Stahl, D.A., 2018. Brock Biology of Microorganisms, 15th Global edition. Boston, US: Benjamin Cummins.

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), pp.3141-3155.

Mainali, K.P., Bewick, S., Thielen, P., Mehoke, T., Breitwieser, F.P., Paudel, S., Adhikari, A., Wolfe, J., Slud, E.V., Karig, D. and Fagan, W.F., 2017. Statistical analysis of co-occurrence patterns in microbial presence-absence datasets. PLoS One, 12(11), p.e0187132.

Martens, E.C., Koropatkin, N.M., Smith, T.J. and Gordon, J.I., 2009. Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. Journal of Biological Chemistry, 284(37), pp.24673-24677.

Martin, C., Binda, C., Fraaije, M.W. and Mattevi, A., 2020. The multipurpose family of flavoprotein oxidases. In The Enzymes (Vol. 47, pp. 63-86). Academic Press.

Mathai, P.P., Zitomer, D.H. and Maki, J.S., 2015. Quantitative detection of syntrophic fatty acid-degrading bacterial communities in methanogenic environments. Microbiology, 161(6), pp.1189-1197.

Matthies, C. and Schink, B., 1992. Reciprocal isomerization of butyrate and isobutyrate by the strictly anaerobic bacterium strain WoG13 and methanogenic isobutyrate degradation by a defined triculture. Applied and environmental microbiology, 58(5), pp.1435-1439.

Mauky, E., Jacobi, H.F., Liebetrau, J. and Nelles, M., 2015. Flexible biogas production for demand-driven energy supply–feeding strategies and types of substrates. Bioresource technology, 178, pp.262-269.

McBee, R.H., 1954. The characteristics of Clostridium thermocellum. Journal of bacteriology, 67(4), p.505.

McInerney, M.J., Bryant, M.P., Hespell, R.B., Costerton, J.W., 1981. Syntrophomonas wolfei gen. nov. sp. nov., an Anaerobic, Syntrophic, Fatty Acid-Oxidizing Bacterium. Applied and Environmental Microbiology. 41(4), 1029-1039.

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), p.2224.

Meldal, M., Tornoe, C.W. and Christensen, C., 2002. Peptidotriazoles on Solid Phase:[1, 2, 3]-Triazoles by Regiospecific Copper (I)-Catalyzed 1, 3-Dipolar Cycloaddition of Terminal Alkynes to Azides. J. Org. Chem, 67(9), pp.3057-3064.

Metcalf, L., Eddy, H.P. and Tchobanoglous, G., 1991. Wastewater engineering: treatment, disposal, and reuse (Vol. 4). New York: McGraw-Hill.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A. and Wilkening, J., 2008. The metagenomics RAST server–a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC bioinformatics, 9(1), pp.1-8.

Mikheenko, A., Saveliev, V. and Gurevich, A., 2016. MetaQUAST: evaluation of metagenome assemblies. Bioinformatics, 32(7), pp.1088-1090.

Miller, J.R., Koren, S. and Sutton, G., 2010. Assembly algorithms for next-generation sequencing data. Genomics, 95(6), pp.315-327.

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer, E.L., Tosatto, S.C., Paladin, L., Raj, S., Richardson, L.J. and Finn, R.D., 2021. Pfam: The protein families database in 2021. Nucleic Acids Research, 49(D1), pp.D412-D419.

Miura, Y. and Fulco, A.J., 1975. ω-1, ω-2 and ω-3 hydroxylation of long-chain fatty acids, amides and alcohols by a soluble enzyme system from Bacillus megatyerium. Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism, 388(3), pp.305-317.

Miura, Y., 2013. The biological significance of ω-oxidation of fatty acids. Proceedings of the Japan Academy, Series B, 89(8), pp.370-382.

Moestedt, J., Nordell, E. and Schnürer, A., 2014. Comparison of operating strategies for increased biogas production from thin stillage. Journal of biotechnology, 175, pp.22-30.

Monserrate, E., Leschine, S.B. and Canale-Parola, E., 2001. Clostridium hungatei sp. nov., a mesophilic, N2-fixing cellulolytic bacterium isolated from soil. International journal of systematic and evolutionary microbiology, 51(1), pp.123-132.

Montgomery, H.A.C., Dymock, J.F. and Thom, N.S., 1962. The rapid colorimetric determination of organic acids and their salts in sewage-sludge liquor. Analyst, 87(1041), pp.949-955.

Morozova, O. and Marra, M.A., 2008. Applications of next-generation sequencing technologies in functional genomics. Genomics, 92(5), pp.255-264.

Mosey, F.E. and Fernandes, X.A., 1988. Patterns of hydrogen in biogas from the anaerobic digestion of milk-sugars. In Water Pollution Research and Control Brighton (pp. 187-196). Pergamon.

Mostovenko, E., Deelder, A.M. and Palmblad, M., 2011. Protein expression dynamics during Escherichia coli glucose-lactose diauxie. BMC microbiology, 11(1), pp.1-6.

Mou, X., Sun, S., Edwards, R.A., Hodson, R.E. and Moran, M.A., 2008. Bacterial carbon processing by generalist species in the coastal ocean. Nature, 451(7179), pp.708-711.

Murphy, J.D. and Thamsiriroj, T., 2013. Fundamental science and engineering of the anaerobic digestion process for biogas production. In The biogas handbook (pp. 104-130). Woodhead Publishing.

Musat, N., Halm, H., Winterholler, B., Hoppe, P., Peduzzi, S., Hillion, F., Horreard, F., Amann, R., Jørgensen, B.B. and Kuypers, M.M., 2008. A single-cell view on the ecophysiology of anaerobic phototrophic bacteria. Proceedings of the National Academy of Sciences, 105(46), pp.17861-17866.

Namiki, T., Hachiya, T., Tanaka, H. and Sakakibara, Y., 2012. MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads. Nucleic acids research, 40(20), pp.e155-e155.

Narihiro, T., Nobu, M.K., Kim, N.K., Kamagata, Y. and Liu, W.T., 2015. The nexus of syntrophy-associated microbiota in anaerobic digestion revealed by long-term enrichment and community survey. Environmental microbiology, 17(5), pp.1707-1720.

Narihiro, T., Nobu, M.K., Tamaki, H., Kamagata, Y. and Liu, W.T., 2016. Draft genome sequence of Syntrophomonas wolfei subsp. methylbutyratica strain 4J5T (JCM 14075), a mesophilic butyrate-and 2-methylbutyrate-degrading syntroph. Genome Announc., 4(2), pp.e00047-16.

Náthia-Neves, G., Berni, M., Dragone, G., Mussatto, S.I. and Forster-Carneiro, T., 2018. Anaerobic digestion process: technological aspects and recent developments. International journal of environmental science and technology, 15(9), pp.2033-2046.

Nesvizhskii, A.I., Keller, A., Kolker, E. and Aebersold, R., 2003. A statistical model for identifying proteins by tandem mass spectrometry. Analytical chemistry, 75(17), pp.4646-4658.

Neufeld, J.D., Chen, Y., Dumont, M.G. and Murrell, J.C., 2008. Marine methylotrophs revealed by stable-isotope probing, multiple displacement amplification and metagenomics. Environmental microbiology, 10(6), pp.1526-1535.

Neufeld, J.D., Vohra, J., Dumont, M.G., Lueders, T., Manefield, M., Friedrich, M.W. and Murrell, J.C., 2007. DNA stable-isotope probing. Nature protocols, 2(4), pp.860-866.

Neves, L., Pereira, M.A., Mota, M. and Alves, M.M., 2009. Detection and quantification of long chain fatty acids in liquid and solid samples and its relevance to understand anaerobic digestion of lipids. Bioresource Technology, 100(1), pp.91-96.

Nguyen, P.H.L., Kuruparan, P., Visvanathan, C., 2007. Anaerobic digestion of municipal solid waste as a treatment prior to landfill. Bioresource Technology. 98(2), 380-387 https://doi.org/10.1016/j.biortech.2005.12.018

Nicholls, H., 2015. Determining the Microbial Community Dynamics of Anaerobic Digestion Using Metagenomics (Doctoral dissertation, University of York).

Nielsen, L.P., Risgaard-Petersen, N., Fossing, H., Christensen, P.B. and Sayama, M., 2010. Electric currents couple spatially separated biogeochemical processes in marine sediment. Nature, 463(7284), pp.1071-1074.

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, p.115955.

Nishiyama, T., Ueki, A., Kaku, N., Watanabe, K. and Ueki, K., 2009. Bacteroides graminisolvens sp. nov., a xylanolytic anaerobe isolated from a methanogenic reactor treating cattle waste. International journal of systematic and evolutionary microbiology, 59(8), pp.1901-1907.

Nurk, S., Meleshko, D., Korobeynikov, A. and Pevzner, P.A., 2017. metaSPAdes: a new versatile metagenomic assembler. Genome research, 27(5), pp.824-834.

Odnell, A., Recktenwald, M., Stensén, K., Jonsson, B.H. and Karlsson, M., 2016. Activity, life time and effect of hydrolytic enzymes for enhanced biogas production from sludge anaerobic digestion. Water research, 103, pp.462-471.

Oren, A., 2014. The family methanosarcinaceae. The Prokaryotes. Berlin: Springer, pp.259-81.

Orphan, V.J., Turk, K.A., Green, A.M. and House, C.H., 2009. Patterns of 15N assimilation and growth of methanotrophic ANME-2 archaea and sulfate-reducing bacteria within structured syntrophic consortia revealed by FISH-SIMS. Environmental microbiology, 11(7), pp.1777-1791.

O'Sullivan, C.A., Burrell, P.C., Clarke, W.P. and Blackall, L.L., 2005. Structure of a cellulose degrading bacterial community during anaerobic digestion. Biotechnology and Bioengineering, 92(7), pp.871-878.

Parawira, W., Murto, M., Read, J.S. and Mattiasson, B., 2005. Profile of hydrolases and biogas production during two-stage mesophilic anaerobic digestion of solid potato waste. Process Biochemistry, 40(9), pp.2945-2952.

Park, Y.K., Ledesma-Amaro, R. and Nicaud, J.M., 2020. De novo biosynthesis of odd-chain fatty acids in Yarrowia lipolytica enabled by modular pathway engineering. Frontiers in bioengineering and biotechnology, 7, p.484.

Parks, D.H., Chuvochina, M., Rinke, C., Mussig, A.J., Chaumeil, P.A. and Hugenholtz, P., 2021. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. Nucleic Acids Res.

Pavlostathis, S.G. and Giraldo-Gomez, E., 1991. Kinetics of anaerobic treatment: a critical review. Critical Reviews in Environmental Science and Technology, 21(5-6), pp.411-490.

Peng, Y., Leung, H.C., Yiu, S.M. and Chin, F.Y., 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics, 28(11), pp.1420-1428.

Petriglieri, F., Nierychlo, M., Nielsen, P.H. and McIlroy, S.J., 2018. In situ visualisation of the abundant Chloroflexi populations in full-scale anaerobic digesters and the fate of immigrating species. PloS one, 13(11), p.e0206255.

Pezzi, H.M., Niles, D.J., Schehr, J.L., Beebe, D.J. and Lang, J.M., 2018. Integration of magnetic bead-based cell selection into complex isolations. ACS omega, 3(4), pp.3908-3917.

Pholwat, S., Liu, J., Taniuchi, M., Chinli, R., Pongpan, T., Thaipisutikul, I., Ratanakorn, P., Platts-Mills, J.A., Fleece, M., Stroup, S. and Gratz, J., 2019. Genotypic antimicrobial resistance assays for use on E. coli isolates and stool specimens. PloS one, 14(5), p.e0216747.

Playne, M.J., 1985. Determination of ethanol, volatile fatty acids, lactic and succinic acids in fermentation liquids by gas chromatography. Journal of the Science of Food and Agriculture, 36(8), pp.638-644.

Plugge, C.M., 2017. Biogas. Microbial Biotechnology 10(5), 1128-1130.

Poilane, I., Karjalainen, T., Barc, M.C., Bourlioux, P. and Collignon, A., 1998. Protease activity of Clostridium difficile strains. Canadian journal of microbiology, 44(2), pp.157-161.

Ponpium, P., Ratanakhanokchai, K. and Kyu, K.L., 2000. Isolation and properties of a cellulosome-type multienzyme complex of the thermophilic Bacteroides sp. strain P-1. Enzyme and microbial technology, 26(5-6), pp.459-465.

Pontes, H., Guedes de Pinho, P., Casal, S., Carmo, H., Santos, A., Magalhães, T., Remião, F., Carvalho, F. and Bastos, M.L., 2009. GC determination of acetone, acetaldehyde, ethanol, and methanol in biological matrices and cell culture. Journal of chromatographic science, 47(4), pp.272-278.

Pore, S.D., Shetty, D., Arora, P., Maheshwari, S. and Dhakephalkar, P.K., 2016. Metagenome changes in the biogas producing community during anaerobic digestion of rice straw. Bioresource technology, 213, pp.50-53.

Pöschl, M., Ward, S. and Owende, P., 2010. Evaluation of energy efficiency of various biogas production and utilization pathways. Applied energy, 87(11), pp.3305-3321.

Pratscher, J., Dumont, M.G. and Conrad, R., 2011. Ammonia oxidation coupled to CO2 fixation by archaea and bacteria in an agricultural soil. Proceedings of the National Academy of Sciences, 108(10), pp.4170-4175.

Radajewski, S., Ineson, P., Parekh, N.R. and Murrell, J.C., 2000. Stable-isotope probing as a tool in microbial ecology. Nature, 403(6770), pp.646-649.

Radajewski, S., McDonald, I.R. and Murrell, J.C., 2003. Stable-isotope probing of nucleic acids: a window to the function of uncultured microorganisms. Current opinion in biotechnology, 14(3), pp.296-302.

Raposo, F., Borja, R., Martín, M.A., Martín, A., De la Rubia, M.A. and Rincón, B., 2009. Influence of inoculum–substrate ratio on the anaerobic digestion of sunflower oil cake in batch mode: process stability and kinetic evaluation. Chemical Engineering Journal, 149(1-3), pp.70-77.

Rasi, S., Veijanen, A. and Rintala, J., 2007. Trace compounds of biogas from different biogas production plants. Energy, 32(8), pp.1375-1380.

Rasmussen, R., 2001. Quantification on the LightCycler. In Rapid cycle real-time PCR (pp. 21-34). Springer, Berlin, Heidelberg.

Razaviarani, V. and Buchanan, I.D., 2014. Reactor performance and microbial community dynamics during anaerobic co-digestion of municipal wastewater sludge with restaurant grease waste at steady state and overloading stages. Bioresource technology, 172, pp.232-240.

Reichart, N.J., Jay, Z.J., Krukenberg, V., Parker, A.E., Spietz, R.L. and Hatzenpichler, R., 2020. Activity-based cell sorting reveals responses of uncultured archaea and bacteria to substrate amendment. The ISME Journal, 14(11), pp.2851-2861.

Reintjes, G., Arnosti, C., Fuchs, B. and Amann, R., 2019. Selfish, sharing and scavenging bacteria in the Atlantic Ocean: a biogeographical study of bacterial substrate utilisation. The ISME journal, 13(5), pp.1119-1132.

Reintjes, G., Arnosti, C., Fuchs, B.M. and Amann, R., 2017. An alternative polysaccharide uptake mechanism of marine bacteria. The ISME journal, 11(7), pp.1640-1650.

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.F., Darling, A., Malfatti, S., Swan, B.K., Gies, E.A. and Dodsworth, J.A., 2013. Insights into the phylogeny and coding potential of microbial dark matter. Nature, 499(7459), pp.431-437.

Rinzema, A., Boone, M., van Knippenberg, K. and Lettinga, G., 1994. Bactericidal effect of long chain fatty acids in anaerobic digestion. Water Environment Research, 66(1), pp.40-49.

Rittershaus, E.S., Baek, S.H. and Sassetti, C.M., 2013. The normalcy of dormancy: common themes in microbial quiescence. Cell host & microbe, 13(6), pp.643-651.

Riviere, D., Desvignes, V., Pelletier, E., Chaussonnerie, S., Guermazi, S., Weissenbach, J., Li, T., Camacho, P. and Sghir, A., 2009. Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. The ISME journal, 3(6), pp.700-714.

Rocamora, I., Wagland, S.T., Villa, R., Simpson, E.W., Fernández, O. and Bajón-Fernández, Y., 2020. Dry anaerobic digestion of organic waste: A review of operational parameters and their impact on process performance. Bioresource technology, 299, p.122681.

Rodríguez-Moyá, M. and Gonzalez, R., 2015. Proteomic analysis of the response of Escherichia coli to short-chain fatty acids. Journal of proteomics, 122, pp.86-99.

Rostovtsev, V.V., Green, L.G., Fokin, V.V. and Sharpless, K.B., 2002. A stepwise huisgen cycloaddition process: copper (I)-catalyzed regioselective "ligation" of azides and terminal alkynes. Angewandte Chemie, 114(14), pp.2708-2711.

Rotaru, A.E., Shrestha, P.M., Liu, F., Markovaite, B., Chen, S., Nevin, K.P. and Lovley, D.R., 2014b. Direct interspecies electron transfer between Geobacter metallireducens and Methanosarcina barkeri. Applied and environmental microbiology, 80(15), pp.4599-4605.

Rotaru, A.E., Shrestha, P.M., Liu, F., Shrestha, M., Shrestha, D., Embree, M., Zengler, K., Wardman, C., Nevin, K.P. and Lovley, D.R., 2014a. A new model for electron flow during anaerobic digestion: direct interspecies electron transfer to Methanosaeta for the reduction of carbon dioxide to methane. Energy & Environmental Science, 7(1), pp.408-415.

Roy, F., Samain, E., Dubourguier, H.C., Albagnac, G., 1986. Synthrophomonas sapovorans sp. nov., a new obligately proton reducing anaerobe oxidizing

saturated and unsaturated long chain fatty acids. Arch. Microbiol. 145, 142-147.

Sabathé, F., Bélaïch, A. and Soucaille, P., 2002. Characterization of the cellulolytic complex (cellulosome) of Clostridium acetobutylicum. FEMS microbiology letters, 217(1), pp.15-22.

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.

Sakka, M., Goto, M., Fujino, T., Fujino, E., Karita, S., Kimura, T. and Sakka, K., 2010. Analysis of a Clostridium josui cellulase gene cluster containing the man5A gene and characterization of recombinant Man5A. Bioscience, biotechnology, and biochemistry, 74(10), pp.2077-2082.

Samo, T.J., Smriga, S., Malfatti, F., Sherwood, B.P. and Azam, F., 2014. Broad distribution and high proportion of protein synthesis active marine bacteria revealed by click chemistry at the single cell level. Frontiers in Marine Science, 1, p.48.

Sawers, R.G., Fischer, M. and Falke, D., 2019. Anaerobic nitrate respiration in the aerobe Streptomyces coelicolor A3 (2): helping maintain a proton gradient during dormancy. Environmental microbiology reports, 11(5), pp.645-650.

Schatz, M.C., Delcher, A.L. and Salzberg, S.L., 2010. Assembly of large genomes using second-generation sequencing. Genome research, 20(9), pp.1165-1173.

Schink, B., 1985. Mechanisms and kinetics of succinate and propionate degradation in anoxic freshwater sediments and sewage sludge. Microbiology, 131(3), pp.643-650.

Schink, B., 1997. Energetics of syntrophic cooperation in methanogenic degradation. Microbiol Mol Biol Rev 61, 262-280.

Schloss, P.D. and Handelsman, J., 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl. Environ. Microbiol., 71(3), pp.1501-1506.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J. and Sahl, J.W., 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and environmental microbiology, 75(23), pp.7537-7541.

Schlüter, A., Bekel, T., Diaz, N.N., Dondrup, M., Eichenlaub, R., Gartemann, K.H., Krahn, I., Krause, L., Krömeke, H., Kruse, O. and Mussgnug, J.H., 2008. The metagenome of a biogas-producing microbial community of a production-scale biogas plant fermenter analysed by the 454-pyrosequencing technology. Journal of biotechnology, 136(1-2), pp.77-90.

Scholten-Koerselman, I., Houwaard, F., Janssen, P. and Zehnder, A.J., 1986. Bacteroides xylanolyticus sp. nov., a xylanolytic bacterium from methane producing cattle manure. Antonie Van Leeuwenhoek, 52(6), pp.543-554.

Schönfeld, P. and Wojtczak, L., 2016. Short-and medium-chain fatty acids in energy metabolism: the cellular perspective. Journal of lipid research, 57(6), pp.943-954.

Schulz, H., 1991. Beta oxidation of fatty acids. Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism, 1081(2), pp.109-120.

Schwarz, W., 2001. The cellulosome and cellulose degradation by anaerobic bacteria. Applied microbiology and biotechnology, 56(5-6), pp.634-649.

Schwarz, W.H., Bronnenmeier, K., Landmann, B., Wanner, G., Staudenbauer, W.L., Kurose, N. and Takayama, T., 1995. Molecular characterization of four strains of the cellulolytic thermophile Clostridium stercorarium. Bioscience, biotechnology, and biochemistry, 59(9), pp.1661-1665.

Sebaihia, M., Peck, M.W., Minton, N.P., Thomson, N.R., Holden, M.T., Mitchell, W.J., Carter, A.T., Bentley, S.D., Mason, D.R., Crossman, L. and Paul, C.J., 2007. Genome sequence of a proteolytic (Group I) Clostridium botulinum strain Hall A and comparative analysis of the clostridial genomes. Genome research, 17(7), pp.1082-1092.

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics, 30(14), pp.2068-2069.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S. and Huttenhower, C., 2011. Metagenomic biomarker discovery and explanation. Genome biology, 12(6), pp.1-18.

Sekiguchi, Y., Kamagata, Y., Nakamura, K., Ohashi, A., Harada, H., 2000. Syntrophothermus lipocalidus gen. nov., sp. nov., a novel thermophilic, syntrophic, fatty-acid-oxidizing anaerobe which utilizes isobutyrate. International Journal of Systematic and Evolutionary Microbiology. 50, 771-779.

Seppänen-Laakso, T., Laakso, I. and Hiltunen, R., 2002. Analysis of fatty acids by gas chromatography, and its relevance to research on health and nutrition. Analytica Chimica Acta, 465(1-2), pp.39-62.

Shafin, K., Pesout, T., Lorig-Roach, R., Haukness, M., Olsen, H.E., Bosworth, C., Armstrong, J., Tigyi, K., Maurer, N., Koren, S. and Sedlazeck, F.J., 2020. Nanopore sequencing and the Shasta toolkit enable efficient de novo assembly of eleven human genomes. Nature biotechnology, 38(9), pp.1044-1053.

Shiratori, H., Sasaya, K., Ohiwa, H., Ikeno, H., Ayame, S., Kataoka, N., Miya, A., Beppu, T. and Ueda, K., 2009. Clostridium clariflavum sp. nov. and Clostridium caenicola sp. nov., moderately thermophilic, cellulose-/cellobiose-digesting bacteria isolated from methanogenic sludge.

International journal of systematic and evolutionary microbiology, 59(7), pp.1764-1770.

Shoham, Y., Lamed, R. and Bayer, E.A., 1999. The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides. Trends in microbiology, 7(7), pp.275-281.

Siedlecka, E.M., Kumirska, J., Ossowski, T., Glamowski, P., Gołębiowski, M., Gajdus, J., Kaczyński, Z. and Stepnowski, P., 2008. Determination of volatile fatty acids in environmental aqueous samples. Polish Journal of Environmental Studies, 17(3), pp.351-356.

Singer, E., Wagner, M. and Woyke, T., 2017. Capturing the genetic makeup of the active microbiome in situ. The ISME journal, 11(9), pp.1949-1963.

Soares, L.A., Rabelo, C.A.B.S., Sakamoto, I.K., Delforno, T.P., Silva, E.L. and Varesche, M.B.A., 2018. Metagenomic analysis and optimization of hydrogen production from sugarcane bagasse. Biomass and Bioenergy, 117, pp.78-85.

Sohaskey, C.D. and Wayne, L.G., 2003. Role of narK2X and narGHJI in Hypoxic Upregulation of Nitrate Reduction by Mycobacterium tuberculosis. Journal of bacteriology, 185(24), pp.7247-7256.

Sousa, D. Z., Pereira, M. A., Stams, A. J., Alves, M. M. and Smidt, H., 2007b. Microbial communities involved in anaerobic degradation of unsaturated or saturated long-chain fatty acids. Appl Environ Microbiol. 73, 1054–1064.

Sousa, D.Z., Smidt, H., Alves, M.M. and Stams, A.J., 2007a. Syntrophomonas zehnderi sp. nov., an anaerobe that degrades long-chain fatty acids in co-culture with Methanobacterium formicicum. Int J Syst Evol Microbiol 57, 609-615.

Sousa, D.Z., Smidt, H., Alves, M.M., Stams, A.J.M., 2009. Ecophysiology of syntrophic communities that degrade saturated and unsaturated long-chain fatty acids. FEMS Microbiol Ecol. 68, 257-272.

Speece, R.E., 1983. Anaerobic biotechnology for industrial wastewater treatment. Environmental science & technology, 17(9), pp.416A-427A.

Speirs, L., Rice, D.T., Petrovski, S. and Seviour, R.J., 2019. The phylogeny, biodiversity, and ecology of the chloroflexi in activated sludge. Frontiers in microbiology, 10, p.2015.

Sprinzl, M., 1994. Elongation factor Tu: a regulatory GTPase with an integrated effector. Trends in biochemical sciences, 19(6), pp.245-250.

Sriswasdi, S., Yang, C.C. and Iwasaki, W., 2017. Generalist species drive microbial dispersion and evolution. Nature communications, 8(1), pp.1-8.

Steward, K.F., Eilers, B., Tripet, B., Fuchs, A., Dorle, M., Rawle, R., Soriano, B., Balasubramanian, N., Copié, V., Bothner, B. and Hatzenpichler, R., 2020. Metabolic implications of using bioOrthogonal non-canonical amino acid

tagging (BONCAT) for tracking protein synthesis. Frontiers in microbiology, 11, p.197.

Stieb, M. and Schink, B., 1985. Anaerobic oxidation of fatty acids by Clostridium bryantii sp. nov., a sporeforming, obligately syntrophic bacterium. Archives of microbiology, 140(4), pp.387-390.

Stieb, M. and Schink, B., 1986. Anaerobic degradation of isovalerate by a defined methanogenic coculture. Archives of microbiology, 144(3), pp.291-295.

Stieb, M. and Schink, B., 1989. Anaerobic degradation of isobutyrate by methanogenic enrichment cultures and by a Desulfococcus multivorans strain. Archives of microbiology, 151(2), pp.126-132.

Sträuber, H., Schröder, M. and Kleinsteuber, S., 2012. Metabolic and microbial community dynamics during the hydrolytic and acidogenic fermentation in a leach-bed process. Energy, Sustainability and Society, 2(1), pp.1-10.

Sul, W.J., Park, J., Quensen III, J.F., Rodrigues, J.L., Seliger, L., Tsoi, T.V., Zylstra, G.J. and Tiedje, J.M., 2009. DNA-stable isotope probing integrated with metagenomics for retrieval of biphenyl dioxygenase genes from polychlorinated biphenyl-contaminated river sediment. Applied and Environmental Microbiology, 75(17), pp.5501-5506.

Sun, L., Pope, P.B., Eijsink, V.G. and Schnürer, A., 2015. Characterization of microbial community structure during continuous anaerobic digestion of straw and cow manure. Microbial biotechnology, 8(5), pp.815-827.

Sundberg, C., Al-Soud, W.A., Larsson, M., Alm, E., Yekta, S.S., Svensson, B.H., Sørensen, S.J. and Karlsson, A., 2013. 454 pyrosequencing analyses of bacterial and archaeal richness in 21 full-scale biogas digesters. FEMS microbiology ecology, 85(3), pp.612-626.

Svetlitshnyi, V., Rainey, F., Wiegel, J., 1996. Thermosyntropha lipolytica gen. nov. sp. Nov., a lipolytic, anaerobic, alkalitolerant, thermophilic bacterium utilizing short- and long-chain fatty acids in syntrophic coculture with a methanogenic archaeum. International Journal of Systematic Bacteriology. 46(10), 1131-1137.

Tamaru, Y., Karita, S., Ibrahim, A., Chan, H. and Doi, R.H., 2000. A large gene cluster for the Clostridium cellulovorans cellulosome. Journal of bacteriology, 182(20), pp.5906-5910.

Tang, Y.Q., Ji, P., Hayashi, J., Koike, Y., Wu, X.L. and Kida, K., 2011. Characteristic microbial community of a dry thermophilic methanogenic digester: its long-term stability and change with feeding. Applied microbiology and biotechnology, 91(5), pp.1447-1461.

Tao, B., Alessi, A.M., Zhang, Y., Chong, J.P., Heaven, S. and Banks, C.J., 2019. Simultaneous biomethanisation of endogenous and imported CO_2 in organically loaded anaerobic digesters. Applied energy, 247, pp.670-681.

Thabet, O.B.D., Fardeau, M.L., Joulian, C., Thomas, P., Hamdi, M., Garcia, J.L. and Ollivier, B., 2004. Clostridium tunisiense sp. nov., a new proteolytic, sulfur-reducing bacterium isolated from an olive mill wastewater contaminated by phosphogypse. Anaerobe, 10(3), pp.185-190.

Thauer, R.K., Kaster, A.K., Seedorf, H., Buckel, W. and Hedderich, R., 2008. Methanogenic archaea: ecologically relevant differences in energy conservation. Nature Reviews Microbiology, 6(8), pp.579-591.

The Anaerobic Digestion & Bioresources Association (2020), AD Map, ADBA, London, UK. http://adbioresources.org/map

Thierry, A., Richoux, R. and Kerjean, J.R., 2004. Isovaleric acid is mainly produced by Propionibacterium freudenreichii in Swiss cheese. International dairy journal, 14(9), pp.801-807.

Tomei, M.C., Braguglia, C.M., Cento, G. and Mininni, G., 2009. Modeling of anaerobic digestion of sludge. Critical Reviews in Environmental Science and Technology, 39(12), pp.1003-1051.

Tukanghan, W., Hupfauf, S., Gómez-Brandón, M., Insam, H., Salvenmoser, W., Prasertsan, P., Cheirsilp, B. and Sompong, O., 2021. Symbiotic Bacteroides and Clostridium-rich methanogenic consortium enhanced biogas production of high-solid anaerobic digestion systems. Bioresource Technology Reports, 14, p.100685.

Ueki, A., Abe, K., Kaku, N., Watanabe, K. and Ueki, K., 2008. Bacteroides propionicifaciens sp. nov., isolated from rice-straw residue in a methanogenic reactor treating waste from cattle farms. International journal of systematic and evolutionary microbiology, 58(2), pp.346-352.

Ueki, A., Abe, K., Ohtaki, Y., Kaku, N., Watanabe, K. and Ueki, K., 2011. Bacteroides paurosaccharolyticus sp. nov., isolated from a methanogenic reactor treating waste from cattle farms. International journal of systematic and evolutionary microbiology, 61(2), pp.448-453.

Ullmann, I.F., Tunsjø, H.S., Andreassen, M., Nielsen, K.M., Lund, V. and Charnock, C., 2019. Detection of aminoglycoside resistant bacteria in sludge samples from Norwegian drinking water treatment plants. Frontiers in microbiology, 10, p.487.

Urbach, E., Vergin, K.L. and Giovannoni, S.J., 1999. Immunochemical detection and isolation of DNA from metabolically active bacteria. Applied and environmental microbiology, 65(3), pp.1207-1213.

Valentini, T.D., Lucas, S.K., Binder, K.A., Cameron, L.C., Motl, J.A., Dunitz, J.M. and Hunter, R.C., 2020. Bioorthogonal non-canonical amino acid tagging reveals translationally active subpopulations of the cystic fibrosis lung microbiota. Nature communications, 11(1), pp.1-11.

Van Lier, J.B., Grolle, K.C., Frijters, C.T., Stams, A.J. and Lettinga, G., 1993. Effects of acetate, propionate, and butyrate on the thermophilic anaerobic

degradation of propionate by methanogenic sludge and defined cultures. Applied and Environmental Microbiology, 59(4), pp.1003-1011.

Vanwonterghem, I., Jensen, P.D., Ho, D.P., Batstone, D.J. and Tyson, G.W., 2014. Linking microbial community structure, interactions and function in anaerobic digesters using new molecular techniques. Current opinion in biotechnology, 27, pp.55-64.

Varel, V.H., Tanner, R.S. and Woese, C.R., 1995. Clostridium herbivorans sp. nov., a cellulolytic anaerobe from the pig intestine. International Journal of Systematic and Evolutionary Microbiology, 45(3), pp.490-494.

Vasco-Correa, J., Khanal, S., Manandhar, A. and Shah, A., 2018. Anaerobic digestion for bioenergy production: Global status, environmental and techno-economic implications, and government policies. Bioresource technology, 247, pp.1015-1026.

Vaser, R., Sović, I., Nagarajan, N. and Šikić, M., 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome research, 27(5), pp.737-746.

Verastegui, Y., Cheng, J., Engel, K., Kolczynski, D., Mortimer, S., Lavigne, J., Montalibet, J., Romantsov, T., Hall, M., McConkey, B.J. and Rose, D.R., 2014. Multisubstrate isotope labeling and metagenomic analysis of active soil bacterial communities. MBio, 5(4), pp.e01157-14.

Volk, A. and Kähler, C.J., 2018. Density model for aqueous glycerol solutions. Experiments in Fluids, 59(5), pp.1-4.

Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K. and Earl, A.M., 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PloS one, 9(11), p.e112963.

Wallrabenstein, C., Hauschild, E., Schink, B., 1995. Syntrophobacter pfennigii sp. nov., new syntrophically propionate-oxidizing anaerobe growing in pure culture with propionate and sulfate. Archives of Microbiology. 164(5), 346-352.

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), pp.214-222.

Walters, K.E. and Martiny, J.B., 2020. Alpha-, beta-, and gamma-diversity of bacteria varies across habitats. Plos one, 15(9), p.e0233872.

Wan, C., Zhou, Q., Fu, G. and Li, Y., 2011. Semi-continuous anaerobic co-digestion of thickened waste activated sludge and fat, oil and grease. Waste management, 31(8), pp.1752-1758.

Wang, L., Aziz, T.N. and Francis, L., 2013. Determining the limits of anaerobic co-digestion of thickened waste activated sludge with grease interceptor waste. Water research, 47(11), pp.3835-3844.

Wang, Q., Kuninobu, M., Ogawa, H.I. and Kato, Y., 1999. Degradation of volatile fatty acids in highly efficient anaerobic digestion. Biomass and Bioenergy, 16(6), pp.407-416.

Wang, Z., Zhang, X.X., Huang, K., Miao, Y., Shi, P., Liu, B., Long, C. and Li, A., 2013. Metagenomic profiling of antibiotic resistance genes and mobile genetic elements in a tannery wastewater treatment plant. PloS one, 8(10), p.e76079.

Weissman, J.L., Hou, S. and Fuhrman, J.A., 2021. Estimating maximal microbial growth rates from cultures, metagenomes, and single cells via codon usage patterns. Proceedings of the National Academy of Sciences, 118(12).

Wells, C.L. and Wilkins, T.D., 1996. Clostridia: sporeforming anaerobic bacilli. In Medical Microbiology. 4th edition. University of Texas Medical Branch at Galveston.

Welte, C. and Deppenmeier, U., 2014. Bioenergetics and anaerobic respiratory chains of aceticlastic methanogens. Biochimica et Biophysica Acta (BBA)-Bioenergetics, 1837(7), pp.1130-1147.

Wemheuer, F., Taylor, J.A., Daniel, R., Johnston, E., Meinicke, P., Thomas, T. and Wemheuer, B., 2020. Tax4Fun2: prediction of habitat-specific functional profiles and functional redundancy based on 16S rRNA gene sequences. Environmental Microbiome, 15, pp.1-12.

Werner, J.J., Garcia, M.L., Perkins, S.D., Yarasheski, K.E., Smith, S.R., Muegge, B.D., Stadermann, F.J., DeRito, C.M., Floss, C., Madsen, E.L. and Gordon, J.I., 2014. Microbial community dynamics and stability during an ammonia-induced shift to syntrophic acetate oxidation. Applied and environmental microbiology, 80(11), pp.3375-3383.

Werner, J.J., Knights, D., Garcia, M.L., Scalfone, N.B., Smith, S., Yarasheski, K., Cummings, T.A., Beers, A.R., Knight, R. and Angenent, L.T., 2011. Bacterial community structures are unique and resilient in full-scale bioenergy systems. Proceedings of the National Academy of Sciences, 108(10), pp.4158-4163.

Whitehead, T.R., Cotta, M.A., Collins, M.D., Falsen, E. and Lawson, P.A., 2005. Bacteroides coprosuis sp. nov., isolated from swine-manure storage pits. International journal of systematic and evolutionary microbiology, 55(6), pp.2515-2518.

Whiteley, A.S., Thomson, B., Lueders, T. and Manefield, M., 2007. RNA stable-isotope probing. Nature Protocols, 2(4), p.838.

Whitman, W.B., Boone, D.R., Koga, Y., Keswani, J., 2001. Taxonomy of methanogenic archaea. In: Boone DR, Castenholtz RW, Garrity GM (eds) Bergey's manual of systematic bacteriology, vol 1, 2nd edn. Springer, New York, pp 211–213.

Wilén, B.M., Liébana, R., Persson, F., Modin, O. and Hermansson, M., 2018. The mechanisms of granulation of activated sludge in wastewater treatment, its optimization, and impact on effluent quality. Applied microbiology and biotechnology, 102(12), pp.5005-5020.

Wilkins, D., Lu, X.Y., Shen, Z., Chen, J. and Lee, P.K., 2015. Pyrosequencing of mcrA and archaeal 16S rRNA genes reveals diversity and substrate preferences of methanogen communities in anaerobic digesters. Appl. Environ. Microbiol., 81(2), pp.604-613.

Winderl, C., Penning, H., Von Netzer, F., Meckenstock, R.U. and Lueders, T., 2010. DNA-SIP identifies sulfate-reducing Clostridia as important toluene degraders in tar-oil-contaminated aquifer sediment. The ISME Journal, 4(10), pp.1314-1325.

Wirth, R., Kovács, E., Maróti, G., Bagi, Z., Rákhely, G. and Kovács, K.L., 2012. Characterization of a biogas-producing microbial community by short-read next generation DNA sequencing. Biotechnology for biofuels, 5(1), pp.1-16.

Woese, C.R. and Fox, G.E., 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proceedings of the National Academy of Sciences, 74(11), pp.5088-5090.

Wong, M.T., Zhang, D., Li, J., Hui, R.K.H., Tun, H.M., Brar, M.S., Park, T.J., Chen, Y. and Leung, F.C., 2013. Towards a metagenomic understanding on enhanced biomethane production from waste activated sludge after pH 10 pretreatment. Biotechnology for Biofuels, 6(1), p.38.

Wood, D.E. and Salzberg, S.L., 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome biology, 15(3), pp.1-12.

Worm, P., Koehorst, J.J., Visser, M., Sedano-Núñez, V.T., Schaap, P.J., Plugge, C.M., Sousa, D.Z. and Stams, A.J., 2014. A genomic view on syntrophic versus non-syntrophic lifestyle in anaerobic fatty acid degrading communities. Biochimica et Biophysica Acta (BBA)-Bioenergetics, 1837(12), pp.2004-2016.

Wu, C., Dong, X., Liu, X., 2007. Syntrophomonas wolfei subsp. methylbutyratica subsp. nov., and assignment of Syntrophomonas wolfei subsp. saponavida to Syntrophomonas saponavida sp. nov. comb. nov. Systematic and Applied Microbiology. 30, 376-380.

Wu, C., Liu, X., Dong, X., 2006a. Syntrophomonas cellicola sp. nov., a sporeforming syntrophic bacterium isolated from a distilled-spirit-fermenting cellar, and assignment of Syntrophospora bryantii to Syntrophomonas bryantii comb. nov. International Journal of Systematic and Evolutionary Microbiology. 56, 2331-2335.

Wu, C., Liu, X., Dong, X., 2006b. Syntrophomonas erecta subsp. sporosyntropha subsp. nov., a spore-forming bacterium that degrades short chain fatty acids in co-culture with methanogens. Systematic and Applied Microbiology. 29, 457-462.

Wu, W.M., Jain, M.K. and Zeikus, J.G., 1994. Anaerobic degradation of normal- and branched-chain fatty acids with four or more carbons to methane by a syntrophic methanogenic triculture. Applied and environmental microbiology, 60(7), pp.2220-2226.

Wu, Y.R. and He, J., 2013. Characterization of anaerobic consortia coupled lignin depolymerization with biomethane generation. Bioresource technology, 139, pp.5-12.

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), pp.605-607.

Xin, F., Dong, W., Zhang, W., Ma, J. and Jiang, M., 2019. Biobutanol production from crystalline cellulose through consolidated bioprocessing. Trends in biotechnology, 37(2), pp.167-180.

Yang, J.C., Chynoweth, D.P., Williams, D.S. and Li, A., 1990. Clostridium aldrichii sp. nov., a cellulolytic mesophile inhabiting a wood-fermenting anaerobic digester. International Journal of Systematic and Evolutionary Microbiology, 40(3), pp.268-272.

Yang, Y., Yu, K., Xia, Y., Lau, F.T., Tang, D.T., Fung, W.C., Fang, H.H. and Zhang, T., 2014. Metagenomic analysis of sludge from full-scale anaerobic digesters operated in municipal wastewater treatment plants. Applied microbiology and biotechnology, 98(12), pp.5709-5718.

Yanling, H., Youfang, D. and Yanquan, L., 1991. Two cellulolytic Clostridium species: Clostridium cellulosi sp. nov. and Clostridium cellulofermentans sp. nov. International Journal of Systematic and Evolutionary Microbiology, 41(2), pp.306-309.

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.H., Whitman, W.B., Euzéby, J., Amann, R. and Rosselló-Móra, R., 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nature Reviews Microbiology, 12(9), pp.635-645.

Yin, B., Crowley, D., Sparovek, G., De Melo, W.J. and Borneman, J., 2000. Bacterial functional redundancy along a soil reclamation gradient. Applied and environmental microbiology, 66(10), pp.4361-4365.

Yu, D., Kurola, J.M., Lähde, K., Kymäläinen, M., Sinkkonen, A. and Romantschuk, M., 2014. Biogas production and methanogenic archaeal community in mesophilic and thermophilic anaerobic co-digestion processes. Journal of Environmental Management, 143, pp.54-60.

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), pp.248-258.

Zhang, C., Liu, X., Dong, X., 2004. Syntrophomonas curvata sp. nov., an anaerobe that degrades fatty acids in co-culture with methanogens. International Journal of Systematic and Evolutionary Microbiology. 54, 969-973.

Zhang, C., Liu, X., Dong, X., 2005. Syntrophomonas erecta sp. nov., a novel anaerobe that syntrophically degrades short-chain fatty acids. International Journal of Systematic and Evolutionary Microbiology. 55, 799-803.

Zhang, C., Su, H., Baeyens, J. and Tan, T., 2014. Reviewing the anaerobic digestion of food waste for biogas production. Renewable and Sustainable Energy Reviews, 38, pp.383-392.

Zhang, L., Lee, C.H. and Jahng, D., 2011. Restriction of linoleic acid inhibition of methanization of piggery wastewater and enhancement of its mineralization by adding calcium ions. Journal of Chemical Technology & Biotechnology, 86(2), pp.282-289.

Zhao, Z., Li, Y., Quan, X. and Zhang, Y., 2017. Towards engineering application: potential mechanism for enhancing anaerobic digestion of complex organic waste with different types of conductive materials. Water research, 115, pp.266-277.

Zhao, Z., Li, Y., Zhang, Y. and Lovley, D.R., 2020. Sparking Anaerobic Digestion: Promoting Direct Interspecies Electron Transfer to Enhance Methane Production. Iscience, p.101794.

Zhu, J., Zheng, H., Ai, G., Zhang, G., Liu, D., Liu, X. and Dong, X., 2012. The genome characteristics and predicted function of methyl-group oxidation pathway in the obligate aceticlastic methanogens, Methanosaeta spp. PLoS One, 7(5), p.e36756.

Zhu, X., Campanaro, S., Treu, L., Kougias, P.G. and Angelidaki, I., 2019. Novel ecological insights and functional roles during anaerobic digestion of saccharides unveiled by genome-centric metagenomics. Water research, 151, pp.271-279.

Ziels, R.M., Beck, D.A.C., Marti, M., Gough, H.L., Stensel, H.D., Svensson, B.H., 2015. Monitoring the dynamics of syntrophic β-oxidizing bacteria during anaerobic degradation of oleic acid by quantitative PCR. FEMS Microbiology Ecology. 91, 1-13.

Ziels, R.M., Karlsson, A., Beck, D.A., Ejlertsson, J., Yekta, S.S., Bjorn, A., Stensel, H.D. and Svensson, B.H., 2016. Microbial community adaptation influences long-chain fatty acid conversion during anaerobic codigestion of fats, oils, and grease with municipal sludge. Water research, 103, pp.372-382.

Ziels, R.M., Sousa, D.Z., Stensel, H.D. and Beck, D.A., 2018. DNA-SIP based genome-centric metagenomics identifies key long-chain fatty acid-degrading populations in anaerobic digesters with different feeding frequencies. The ISME journal, 12(1), pp.112-123.

Ziels, R.M., Nobu, M.K. and Sousa, D.Z., 2019. Elucidating syntrophic butyrate-degrading populations in anaerobic digesters using stable-isotope-informed genome-resolved metagenomics. Msystems, 4(4), pp.e00159-19.

Zverlov, V.V., Hiegl, W., Köck, D.E., Kellermann, J., Köllmeier, T. and Schwarz, W.H., 2010. Hydrolytic bacteria in mesophilic and thermophilic degradation of plant biomass. Engineering in Life Sciences, 10(6), pp.528-536.