A FUNCTIONAL APPROACH TO CHARACTERISING THE MICROBIAL COMMUNITIES UNDERPINNING LOW-TEMPERATURE ANAEROBIC DIGESTION

A thesis submitted to the National University of Ireland for the degree of Doctor of Philosophy

By:

Eoin Gunnigle

Microbial Ecology Laboratory, Discipline of Microbiology,

School of Natural Sciences,

National University of Ireland, Galway

December 2012

Head of Department/Research Supervisor: Professor Vincent O'Flaherty

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

This research would not have been possible without the help and support of many people.

First and foremost I would like to thank my supervisor, Prof. Vincent O'Flaherty, for providing me with this wonderful opportunity and for his valuable support over the years of my PhD.

I would like to specially thank Dr. Florence Abram, who selflessly made my life easier during my time in the lab through her thoughtful words and constructive guidance.

Thanks also to the rest of the MEL, particularly Niamh, Alma, Paul, Denise and Dermot, who have all contributed to my work in no minor way, and answered many questions I should have known the answers to.

To all the members of the Microbiology Department, staff, students and technicians; it has been a pleasure to work with you all over the past few years, thank you for everything. Particularly, I would like to thank Maurice Martyn for all his help throughout my practical work.

External to NUI Galway, I would like to thanks Dr. Catherine Botting and Dr. Matthew Fuszard in the BSRC Mass Spectrometry and Proteomics Facility at the University of St. Andrews for their brilliant work undertaken with proteomic data. Also, I would like to thank Prof. Jeppe Lund Nielsen at the University of Aalborg for his thoughtful assistance with my MAR-FISH work.

To my parents, Paddy and Theresa, whom I owe a sincere amount of gratitude, especially over the last year, which have been made a lot easier through your patience and support. Thank you.

Many thanks to all my friends micro related and not, you are great.

Abstract

Low-temperature anaerobic digestion (LTAD) presents a sustainable, cost-efficient technology for the treatment of a vast array of wastewater streams. However, this microbially-mediated process requires further understanding and experimental characterisation if large-scale application of LTAD is to be realised. The objective of this thesis was to employ a functional-based approach to characterise the microbial communities underpinning low-temperature anaerobic digestion.

Firstly, the key microbial functional groups present in end-point samples taken from three, laboratory-scale, expanded granular sludge bed (EGSB) bioreactors; R1 (37°C), R2 (15°C) and R3 (7°C) were characterised. Metaproteomics, in conjunction with 16S rRNA gene phylogenetic approaches (clone libraries, qPCR), was applied to record microbial community composition and metaproteomic profiles as a function of bioreactor operating temperature. Clone libraries indicated a predominance of the *Chloroflexi* (21%) and *δ-Proteobacteria* (61%) bacterial groups in R1, with *Firmicutes* (24%) and *Bacteroidetes* (46%) prominent in both R2 and R3. The *Methanosaeta* genus was strongly represented in archaeal clone libraries (29% [R1], 76% [R2] and 91% [R3]). This was reflected in the metaproteomic results with 26 (65%) differentially expressed proteins assigned to this methanogenic group. Also evident from the metaproteomic data were proteins assigned to the bacterial phyla *Proteobacteria*, *Firmicutes* and *Actinobacteria*, while the archaeal orders *Methanobacteriales*, and *Methanomicrobiales* were also represented. Interestingly, the identification of a protein assigned to *Methanosarcina* sp. was not consistent with DNA-based community profiling data, where this methanogenic group was not detected, confirming the importance of employing a functional-based approach in this study.

A pure culture proteomic (iTRAQ, 2-DGE) approach was then employed to uncover the sub-mesophilic functional characteristics of a *Methanosarcina* strain (optimum

growth temperature, 37˚C), with proteins assigned to this group detected in previous LTAD bioreactor trials. New insights into the low-temperature adaptation capacity of this mesophilic methanogen, including differentially expressed proteins during low temperature growth, e.g. elongation factor protein expression during sub-mesophilic adaptation, with high levels of viable cells recorded through this 'adaptation' stage (84% [±9.65% SE; n = 10] of cells were viable after 17 days of growth at 15˚C). Therefore, through this polyphasic approach, the psychrotolerant capacity of this organism was characterised.

Finally, two EGSB bioreactors (R1 & R2) were operated, initially at 37°C with a subsequent temperature drop to 15°C, with biomass samples being taken throughout the trial. PCR-based (clone libraries, qPCR, DGGE [RNA- DNA- derived]) and PCR-independent (specific methanogenic activity [SMA] profiling, microautoradiography fluorescent *in situ* hybridisation [MAR-FISH], and metaproteomics) approaches were employed to investigate the microbial community structure and key functional groups throughout the trial, with particular emphasis on the methanogenic archaea. Once again *Methanosaeata* were prominenet in archaeal community and functional analysis with consistent proteomic profiles recorded between the two LTAD bioreator trials investigated in this study. For example, a bifunctional protein (Mcon_1383) with possible function in riboulose monophosphate (RuMP) pathway was significantly expressed in low temperature biomass for both studies.

Overall, community profiling techniques (clone libraries, qPCR) linked with functional-based (RNA-based DGGE, metaproteomics) approaches employed in this research illustrated the importance and metabolic complexity of *Methanosaeta* in a well functioning LTAD system. Also, the incorporation of functional analysis was justified through the uncovering of discrete community information missed through traditional DNA-based community profiling methods e.g. *Methanospirillum* importance in low temperature biomass during LTAD trial confirmed through metaproteomic and RNA-based DGGE profiles.

Chapter 1
Introduction

This introduction chapter includes relevant information from two peer-reviewed journal articles published during my PhD:

Abram, F., **Gunnigle, E.** and O'Flaherty, V. (2009) Optimisation of protein extraction and 2-DE for metaproteomics of microbial communities from anaerobic wastewater treatment biofilms. *Electro*. **30**(23), 4149-4151.

- Here, I undertook 2-DGE Practical work for this publivation in addition to co-writing the paper with the first author.

Siggins, A., **Gunnigle, E.** and Abram, F. (2012) Exploring mixed microbial community functioning: recent advances in metaproteomics. *FEMS Microbiol*. *Ecol*. **80**, 265–280.

- Here, I contributed to the literature review of certain sections as well as co-writing this paper with the first and second author.

Both papers in full are located in **Appendix 5**

1.1 Life in the Cold

1.1.1 Microbial existence in cold environments

Many habitats on Earth have a low average temperature (<15°C) with more than 80% of the biosphere permanently 'cold' (Simankova *et al*., 2003). These cold environments include the deep ocean (~70% of Earths surface), alpine lakes, Polar Regions, terrestrial environments and the upper atmosphere (Fig. 1.1). Therefore, it is not surprising that microbial life in cold environments is not only evident but flourishing, with the largest portion of the Earths biosphere comprised of organisms that thrive in cold environments (Cavicchioli *et al*., 2002).

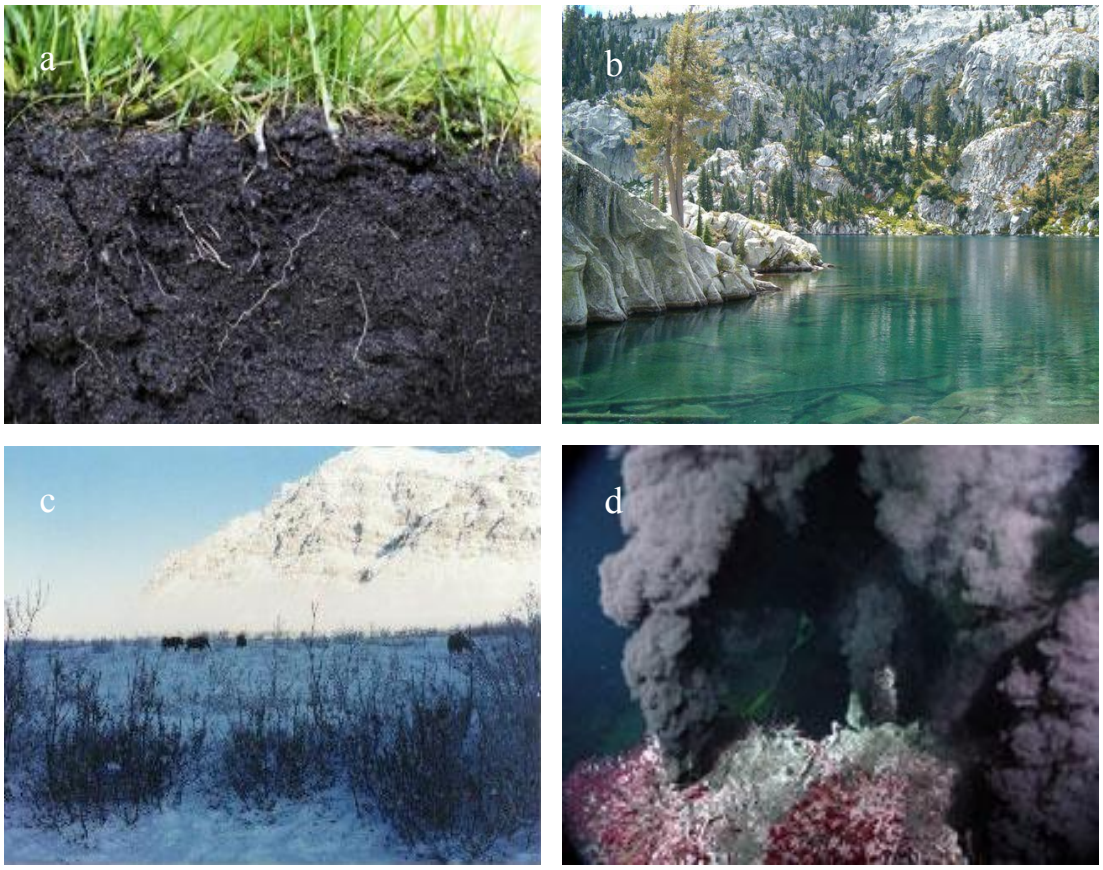


Figure 1.1 Examples of cold environments where microbial communities have been characterised including (a) soil, (b) alpine lakes, (c) tundra and in the vicinity of (d) deep sea hydrothermal vents.

Microbial ecology at temperature extremes presents a key area of research, with the aim of not only providing insights into novel microbial processes and adaptation strategies, but also in the context of environmental change, driven by anthropogenic factors such as global warming. These environments host a wide range of microbial groups with varying thermal requirements and substrate dynamics.

The early categorisation of cold adapted microorganisms was slightly ambiguous due to lack of differentiation between the thermally adapted groups. However, this was rectified with the coining of the term 'psychrophilic' by Morita (1975), which described an organism at thermal equilibrium with their cold environment. Thus, microorganisms displaying a physiological capacity to function at cold temperatures may be classified as being either: (i) psychrophilic- sometimes referred to as the stenopsychrophiles or 'true' psychrophiles, with these organisms having a narrow temperature range for growth of $T_{opt} < 16 - 18°C$ and a maximum temperature for growth of $T_{max} < 25°C$ or (ii) pychrotolerant- capable of growth at a larger temperature range with T_{opt} 18-25°C and $T_{max} > 25°C$ (Cavicchioli, 2006). This latter group can demonstrate a tolerance to temperature fluctuations usually associated with the environmental regime from which the organism was isolated (Russell *et al*., 1990).

The 'cold-loving' psychrophilic and 'cold-adaptable' pschrotolerant organisms are distributed around the globe and contribute significantly to primary production, nutrient cycling and biomass. For example, in the permanenty cold Lake Fryxell (0-15°C) in Antactica, psychrophilic sulfur oxidising bacteria (SOB) closely related to mesophilic *Thiobacillus* species were found to play a key role in the chemical cycling of carbon and sulfur in this system (Sattley & Madigan, 2006). It has been proposed that low-temperature sulfate reduction and methanogenesis is as high in permanently low-temperature sediments as those in temperate systems (Purdy *et al*., 2003). On a cellular level, these microorganisms have demonstrated diverse adaptation mechanisms and strategies to overcome the challenges presented by life in the cold, which are discussed below.

1.1.2 Molecular mechanisms of cold adaptation

Regardless of how microorganisms exhibiting low-temperature growth and function are defined, this microbial group display properties distinct from other thermal classes. The metabolic capacity to grow at low-temperature stems from particular molecular mechanisms, which have been characterised in the numerous physiological, genomic and proteomic experiments carried out on isolates from cold environments (Table 1.1). These studies, in conjunction with experiments undertaken with psychrotolerant organisms, have provided insights into microbial survival and growth strategies employed in low-temperature environments.

It has been proposed that psychrophilic characteristics are relatively conserved amongst the bacterial and archaeal domains (Cavicchioli, 2006). These adaptations include the production of extracellular polymeric substances (EPS). For example, in a study by Junge *et al*. (2004), it was found that the bacterial species *Colwellia psychrerythraea* produced EPS that was proposed to act as a cryoprotectant, which enabled survival within brine-filled ice veins. Also, membrane fluidity alterations (Tarpgaard *et al*., 2006), changes of intracellular solute concentrations (Methe *et al*., 2005) and genetic regulation (Lim *et al*., 2000), have also been reported to contribute to microbial survival in low-temperature environments.

Another key component of low-temperature existence is the synthesis of specific cold-shock and chaperone proteins, which enable genetic regulation and protein synthesis to continue. Although these proteins may overlap between psychrophilic and psychrotolerant organisms, a 'stressed' state may not be apparent in the specialist cold-adapted organism. This is demonstrated in low-temperature studies, such as those undertaken with the psychrophilic methanogen *Methanococcoides burtonii,* which favours protein folding chaperone activity due to the intrinsic flexibility (based on various structural modifications, concise report by D'Amico *et al*., 2006) of the psychrophilic proteins synthesised (Williams *et al*., 2011). This may be in contrast to the cold-induced denaturation-based protein expression undertaken by psychrotolerant organisms.

Table 1.1 Psychrophilic bacterial and archaeal isolates with full genome sequences available

Organism	Taxanomic group	Isolation location	Growth range temp (°C)	Experimental approaches	Reference
Gillisia limnaea	*Bacteroidetes*	Antarctic Lake Fryxell	5-25	Physiological characterisation, Genomics	Van Trappen *et al*. (2004)
Polaribacter irgensii	*Bacteroidetes*	Antarctic sea water	1.5-12	Physiological characterisation, Genomics	Gosink *et al*. (1998)
Polaribacter filamentus	*Bacteroidetes*	Arctic sea water	4-19	Physiological characterisation, Genomics	Gosink *et al*. (1998)
Psychroflexus torquis	*Bacteroidetes*	Antarctic sea ice	10-20	Physiological characterisation, Genomics, Fatty acid profilng	Bowman *et al*. (1998)
Flavobacterium psychrophilum	*Bacteroidetes*	Marine animal intestine	15-18	Physiological characterisation, Genomics, Fatty acid profilng, PAGE	Bernardet *et al*. (1996)
Hymenobacter roseosalivarius	*Bacteroidetes*	Dry Valley soil, Antarctica	10-18	Physiological characterisation, Genomics, Fatty acid profilng,	Hirsch *et al*. (1998)
Octadecabacter antarcticus	*Alphaproteobacteria*	Polar sea ice	4-10	Physiological characterisation, Genomics, Fatty acid profilng,	Gosink *et al*. (1997)

Organism	Taxanomic group	Isolation location	Growth range temp (°C)	Experimental approaches	Reference
Sphingopyxis alaskensis	*Alphaproteobacteria*	Alaskan Bay, Northern sea	4-10	Physiological characterisation, Genomics, Proteomics	Cavicchioli *et al*. (2003)
Pseudoalteromonas haloplanktis	*Gammaproteobacteria*	Antarctic sea water	4-20	Physiological characterisation, Genomics, Fatty acid profiling, Proteomics	Cavicchiolo *et al*. (2003)
Psychromonas boydii	*Gammaproteobacteria*	Arctic polar sea ice	-12-10	Pyhysiological characterisation, Genomics, Fatty acid profiling	Auman *et al*. (2010)
Colwellia psychrerythea	*Gammaproteobacteria*	Arctic marine sediment	2	Pyhysiological characterisation, Genomics,	Deming *et al*. (1998)
Shewanella halifaxensis	*Gammaproteobacteria*	Atlantic marine sediment	4-25	Pyhysiological characterisation, Genomics, Fatty acid profiling	Zhao *et al*. (2006)
Photobacterium profundum	*Gammaproteobacteria*	Deep sea sediment	8-12	Pyhysiological characterisation, Genomics, Fatty acid profiling	Nogi *et al*. (1998)
Oleispira antarctica	*Gammaproteobacteria*	Ross Sea, Antarctica	1-25	Pyhysiological characterisation, Genomics, Fatty acid profiling	Yakimov *et al*. (2003)

Organism	Taxanomic group	Isolation location	Growth range temp (°C)	Experimental approaches	Reference
Psychrobacter articus	*Gammaproteobacteria*	Siberian permafrost	-10-30	Pyhysiological characterisation, Genomics, Fatty acid profiling	Bakermans *et al*. (2006)
Exiguobacterium sibiricum	*Firmicutes*	Siberian permafrost	-2.5-40	Pyhysiological characterisation, Genomics, Fatty acid profiling	Rodrigues *et al*. (2006)
Methanogenium frigidum	*Euryarchaeota*	Ace Lake, Antarctica	1-17	Pyhysiological characterisation, Genomics, Microscopy	Franzmann *et al*. (1997)
Methanococcoides burtonii	*Euryarchaeota*	Ace Lake, Antarctica	-2.5-29.5	Pyhysiological characterisation, Genomics, Microscopy	Franzmann *et al*. (1992)
Crenarchaeum symbiosum	*Crenarchaeota*	Marine sponge tissue	8-18	Pyhysiological characterisation, Genomics	Preston *et al*. (1996)

1.1.3 Microbial communities in low-temperature biotechnology

Microorganisms are ubiquitous in the cold biosphere, with this low-temperature microbial group comprising some of the most metabolically diverse organisms on the planet. Although negative effects on microbial biochemical reactions are evident at low-temperatures, psychrophilic organisms still grow, reproduce and function at similar rates to temperate adapted organisms, thus making them of interest to biotechnological application and research (Feller & Gerday, 2003).

The requirement for all components of a microbial cell to adapt to cold conditions suggests that a wide array of cellular products are available for technological exploitation. Properties, which contribute to the importance of enzymes in low-temperature biotechnology, include high catalytic activity at low-temperatures and low thermostability at elevated temperature (Cavicchioli *et al*., 2002). Examples of important low-temperature enzymes include: pectinase, used in the beverage industry and isolated from *Bacillus* spp. (Cabeza *et al*., 2011); cellulase and xylananse isolated from *Basidiomycetes* spp., with applications in the food industry (Inglis *et al*., 2000); and also proteases purified from the alkaliphilic *Stenotrophomonas maltophilia* (Kuddus & Ramteke, 2009), utilised in psychrophilic detergent applications.

Microbial bioremediation to reduce environmental contamination, for example, after an oil spill event, is an important component of natural waste treatment approaches. Bioremediation relies on the presence of organisms capable of catabolising the organic compound in question (Rogers & McClure, 2003), through optimising endemic populations (biostimulation) or by the introduction of a new species (bioaugmentation). In many incidences, the pollution event occurs in low-temperature environments and, as such, requires a progression of understanding in how best to implement cold-adapted organisms to efficiently treat the polluted site. An example of low-temperature bioremediation activity includes the Antarctic isolate *Pseudomonas* ST41, which demonstrated hydrocarbon reduction at 4°C in oil-contaminated soil from the Signy Islands, Antarctica (Stallwood *et al*., 2005).

The microbial fuel cell (MFC) also represents a potentially viable microbial mediated technology, which has been studied for a number of years but is only recently being proposed as providing a significant opportunity for practical low-temperature applications. MFCs can simultaneously produce a renewable form of energy while treating wastewater (Liu *et al*., 2004). Improvements in design and performance include the modification of electrode surfaces (Logan *et al*., 2007), solution conductivity (Huang and Logan, 2008), and also adapting the microbial species to the particular wastewater being treated (Feng *et al*., 2008). In previous studies, a negative correlation between MFC performance and reduced operating temperature had been observed. For example, a study by Moon *et al*. (2006), found a 21% decrease in power density once a single-chamber MFC had its operating temperature dropped from 35°C to 22°C. However, there has been progression with low-temperature MFC application, with Cheng *et al*. (2011), finding that after applying an initial start up temperature at 30°C with a subsequent drop to as low as 4°C, consistent power generation was achieved. This suggests that once the microbial consortia were formed, they were capable of functioning at sub-mesophilic temperatures.

This low-temperature adaptation phenomenon has also been recorded in other wastewater treatment systems such as laboratory- and industrial- scale bioreactors for the treatment of biodegradable compounds such as simple short-chain fatty acids (VFAs), alcohols and carbohydrates (Siggins *et al*., 2011; O'Reilly *et al*., 2009). The viability of a low-temperature approach to the microbially-mediated treatment of wastewater is investigated below in the context of anaerobic digestion.

1.2 Low-temperature Anaerobic Digestion (LTAD)

1.2.1 The application of LTAD

Anaerobic digestion (AD) involves the biological breakdown of complex organic molecules to biogas, by a consortium of microbes in the absence of oxygen, or alternative electron acceptors, such as nitrate (Saravanane & Murthy, 1999). It is an ancient natural process of decomposition, occurring daily in anoxic environments. Biogas primarily consists of CH_4 (40-60%) and CO_2 (30-50%), which can be utilised as a valuable renewable fuel source.

The biodegradation of organic compounds to methane and CO_2 is accomplished through the sequential and coordinated activity of various microbial (bacterial and archaeal) trophic groups (Fig. 1.2; McCarty & Smith, 1986). The first being hydrolysis, which involves the splitting of a water molecule into a hydrogen ion (H^+) and a hydroxyl ion (OH^-). This is undertaken by facultative and obligate fermentative bacteria. These ions cleave the bond between complex organic polymers, thus reducing them to their monomeric state.

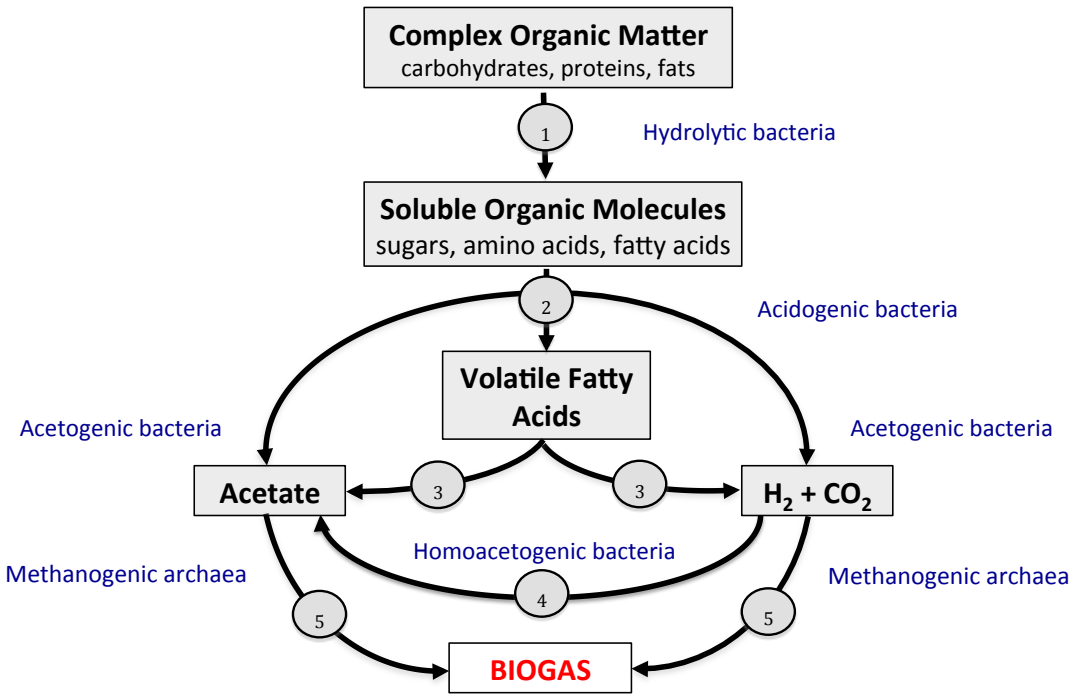


Figure 1.2 The process of Anaerobic Digestion: 1 Hydrolysis; 2 Acidogenesis; 3 Acetogenesis; 4 Homoacetogenesis; 5 Methanogenesis.

The monomeric compounds resulting from hydrolysis are then converted to volatile fatty acids (including acetate, propionate and butyrate), alcohols, ketones and carbon dioxide through the activity of acidogenic bacteria. After this fermentation, acetogenic bacteria convert 3 or 4 carbon intermediates to a mixture of acetate, carbon dioxide and hydrogen. Finally, methanogenic organisms carry out the end-point step using primarily acetate and hydrogen as substrates with the formation of biogas (McCarty & Smith, 1986).

AD treatment is predominantly applied to treat readily biodegradable simplistic wastewaters, such as those produced by agro-food industries (Lettinga *et al*., 2001), with the majority of these AD reactors operated in the mesophilic and thermophilc temperature ranges (30-55°C). The historic prevalence of this regime is based upon microbial growth and reaction rates. However, with the high energy costs associated with mesophilic and high-temperature AD setups (>30% of biogas production used to heat reactors; del Pozo *et al*., 2002), economic feasibility of AD could be enhanced by treating wastewaters at their discharge temperatures. The obvious cost benefit in coordination with an improved understanding of the physical, chemical and biological challenges presented by microbial functioning at low-temperatures (Sotemann *et al*., 2004) has facilitated the progression of low-temperature anaerobic digestion (LTAD) in many successful laboratory-scale studies (Table 1.2).

With respect to the application of LTAD, several pertinent issues were presented, which included reduced substrate utilisation rates (Lin *et al*., 1987), reduced maximum specific growth and methane solubility (Lettinga *et al*., 2001). Nevertheless, successful treatment of an array of wastewater types has been demonstrated at low-temperatures (Table 1.2), which give an indication of the potential for full-scale LTAD application. The development of high-rate reactors, such as the expanded granular sludge bed (EGSB) bioreactor design (Fig. 1.3), with the addition of micro- and ultra- filtration membranes (long solid retention time [SRT]; Martinez-Sosa *et al*., 2011), have played an important role in the successful treatment of wastewaters at low-temperatures in conjunction with the adaptive capabilities of the microbial consortia functioning in these reactors (Table 1.2).

Table 1.2 Examples of successful low-temperature anaerobic digestion studies

Wastewtaer	Temp (°C)	Duration (days)	HRT[c]	CODRE[d]	Biogas CH_4%	Methanogenic community structure	Reference
Glucose	15	120	24-160	83-90	55-77	Acetoclastic methanogenic activity 5-fold higher at end of trial compared with 37°C samples	Akila & Chandra, (2007)
Toluene	15	630	12-48	51-95	57-95	6-fold increase of hydrogenotrophic methanogenic activity in comparison to seed biomass. Reduced acetoclastic activity related to acetate accumulation	Enright *et al*. (2007)
TCE[a]	15-7	609	24	54-86	46-75	Apparent shift from acetate- to hydrogen-mediated methogenesis through qPCR analysis.	Siggins *et al*. (2011)
Molasses	18-5	78	24	25-52	56-77	*Methanomicrobiales* 16S rRNA gene concentration exhibited ca. 32-fold increase from 18°C to 5°C	Zhang *et al*. (2012)
VFA[b]	15-4	1243	12-24	74-93	40-91	Psychrotrophic hydrogenotrophic methanogenesis recorded at end of trial with specific activity ca. 1.5-fold higher at 15°C than at 37°C	McKeown *et al*. (2009)
Glucose	15	300	12-36	52-80	59-67	16S rRNA gene analysis recorded the hydrogenotrophic *Methanocorpusculum* spp. to be important during low-temperature granulation	O'Reilly *et al*. (2010)

[a] Trichloroethylene
[b] Volatile Fatty Acid
[c] Hydraulic Retention Time – all expressed in hours
[d] Chemical Oxygen Demand Removal Efficiency (%)

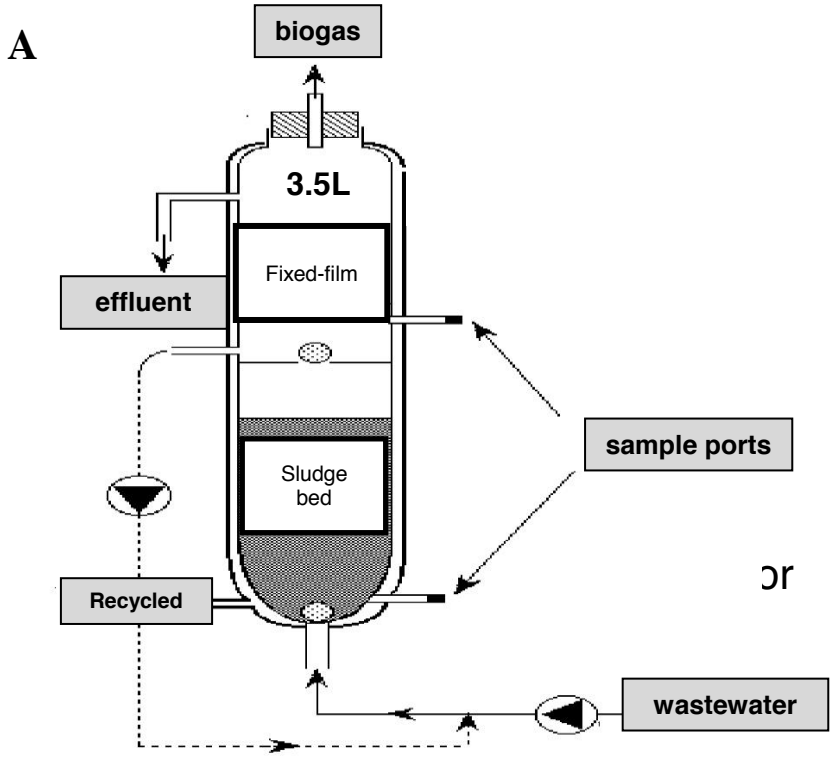


Figure 1.3 (A) lab-scale and (B) pilot scale Expanded Granular Sludge Bed – Anaerobic Filter (EGSB-AF) bioreactor

1.2.2 Microbial dynamics during LTAD

It has been demonstrated that mixed microbial communities functioning in LTAD systems have psychro-active characteristics, regardless if the inoculum used was from a mesophilic reactor (McKeown *et al*., 2012). Therefore, the requirement to source a psychrophilic biomass for successful LTAD wastewater treatment is not necessarily a prerequisite, due to the adaptation potential of the microbial consortia underpinning AD granules. The apparent acclimation of mesophilic biomass to low-temperature conditions within engineered systems may be correlated with successive changes in the community structure (Enright *et al*., 2007; Madden *et al*., 2010).

Whether truly psychrophilic microorganisms prevail in LTAD systems is still to be investigated thoroughly, although it is likely that they may become important in long-term applications, as reported by McKeown *et al*. (2009). Nevertheless, the development of EGSB configuration, with or without the incorporation of an anaerobic filter (AF; Collins *et al*., 2005), or the application of membrane separation approaches (Martinez-Sosa *et al*., 2011), has resulted in efficient retention of the immobilised biomass, promoting the enrichment of efficient methanogenic consortia, through low decay rates (Kd) prevailing at low-temperature (Lettinga *et al*., 2001). The importance of the methanogenic archaea in the environment is discussed below.

1.2.3 The Methanogens

The methanogenic archaea are a diverse group of organisms, representing the most characterised cohort within the archaeal domain (Ferry, 1993). The three primary modes of methanogenic metabolism are (**i**) CO_2 reduction; (**ii**) acetoclastic; and (**iii**) methylotrophic.

(i) $4H_2 + CO_2 \longrightarrow CH_4 + 2H_2O$

$\Delta G°' = -130.4$ kj/mol CH_4 (**1**)

(ii) $CH_3COO^+ + H^+ \longrightarrow CH_4 + CO_2$

$\Delta G°' = -36$ kj/mol CH_4 (**2**)

(iii) $CH_3OH \longrightarrow 3CH_4 + CO_2 + 2H_2O$

$\Delta G°' = -103$ kj/mol CH_4 (**3**)

Many methanogens, particularly members of the *Methanosarcinaceae*, have the metabolic capacity to utilise more than one substrate, for example, *Methanosarcina barkeri* can use H_2 to reduce methanol to methane (Müller *et al*., 1986). The CO_2-reducing methanogens convert carbon from its oxidised low energy state CO_2 into the reduced high energy CH_4, utilising H_2 and generating cellular energy in the process (McCollom, 1999).

The second group utilise the acetate fermentation pathway, using H_2 and acetate. This mode of methanogenesis has been proposed to contribute the highest level of methane produced in the environment (Ferry & Kastead, 2007). This group comprises of only two families, the *Methanosaetaceae* and *Methanosarcinaceae*. The former has been historically categorised as a strict acetoclastic methanogen, although a recent genomic study highlighted the metabolic capacity for possible methyl-group oxidation in three sequenced *Methanosaeta* species (Zhu *et al*., 2012).

This group has a minimum threshold concentration of ~1mM acetate and has been documented to outcompete *Methanosarcinaceae* in environments where acetate concentrations are low (Fey & Conrad, 2000). In addition to acetate, the *Methanosarcinaceae* have the ability to utilise other substrates as stated above, including methylated compounds such as methanol and methylamines, with some species also able to use H_2/CO_2 as a carbon and energy source. Also, the organisms utilising the acetoclastic pathway have been found to benefit from the metabolic functioning of homoacetogenic bacteria (Kotsyurbenko *et al*., 2001). For example,

Beckmann *et al*. (2011), found that in a coal-mine sample, acetoclastic methanogenesis was the preferred methanogenic route even when incubated with isotope labelled H_2-$^{13}CO_2$, with homoacetogenic bacteria responsible for the conversion of hydrogen to acetate.

The third group is the methylotrophic methanogens. This group can utilise waste products from other organisms in the form of simple one-carbon compounds such as methanol and methylamines for methanogenesis. Due to the lack of competition for these specialised substrates, the methylotrophic methanogens play an important role in the global carbon cycle, for example, a study by Guo *et al*. (2012), found that methylotrophic methanogenesis governed biogenic coal bed methane formation at a site in Eastern Ordos Basin, China, which has potential as a renewable energy stream.

Global CH_4 production in low-temperature environments is a pertinent issue, with Erkel *et al*. (2006), suggesting that interactions involving the rice rhizosphere and *in situ* methanogenic species contribute to between 10 – 25% of the annual global CH_4 emissions to the atmosphere. Although mesophilic and thermophilic methanogenesis are comparatively well investigated, research into low-temperature methanogenesis is relatively novel. As such, the microbial pathways and biochemical interactions involved are as yet, poorly understood (Kotsyurbenko *et al*., 1996). There have been varying reports as to whether the hydrogenotrophic or acetoclastic mode of methanogenesis is more prominent at low-temperatures. It is known that acetoclastic methanogenesis may be inhibited in low-temperatures due to acetate accumulation (Kotsyurbenko *et al*., 1996), thus supporting hydrogenotrophic methanogenesis as the preferred route at low-temperature. However, a study by Zepp-Falz *et al*. (1999), suggested that methanogenesis from acetate accounted for the majority of methanogenesis in anoxic sediments. This finding was backed up by a higher number of acetate-utilising methanogens in lake sediments under low-temperature (Lay *et al*., 1998). Thus, the primary mode of methanogenesis in low-temperatures remains unclear, with the environmental regime a possible key factor in determining community structure and key functional groups.

Methanogenic communities are the lynchpin of successful AD systems, regardless of temperature, and thus, understanding the dynamics and community succession of this functional group in a LTAD context is imperative for the successful treatment of wastewater streams. Numerous laboratory-based LTAD studies have investigated methanogenic community structure (Table 1.2), with many of these studies recording an increase in hydrogenotrophic methanogenesis in response to AD at low-temperature. For example, Zhang *et al*. (2012), found a >30-fold increase in the hydrogentrophic *Methanomicrobiales* (16S rRNA gene concentration) when operating temperature was dropped from 18°C to 5°C. Psychrophilic methanogenic characterisation has been limited to a few isolated psychrophilic species (Table 1.1). These include two methanogenic strains isolated from Ace Lake in Antarctica, *Methanococcoides burtonii* and *Methanogenium frigidum*. Together, these methanogens represent the most characterised psychrophilic archaeal species. Although important information has been gathered regarding from these isolates, important functional questions remain relating to *in situ* methanogenic activity at low-temperatures in LTAD systems. Some key knowledge gaps and opportunities for LTAD are presented below.

1.2.4 Knowledge gaps and opportunities for LTAD

Despite the considerable quantity of molecular and micro-analytical information gathered over the past decade, the specific functional roles of individual microorganisms and collective 'community' roles, especially in response to perturbations (e.g. temperature change) are still unresolved (Torsvik *et al*., 2002; Azam & Malfatti, 2007). As such, the advancement of research into the structure, function and biological properties of the microbial communities underpinning LTAD will contribute towards the future improvement and optimisation of LTAD technology.

Anaerobic bioreactors have historically been designed and operated as 'black box' entities, with material balances including, for example, methane content and effluent quality, being the principal tools employed to monitor the AD process. However, this

approach is insufficient in the context of optimal process design and performance with key parameters, such as the routes of carbon flow and the identification of key trophic groups being required, which may in turn give an improved understanding of the key AD microbially-mediated processes, such as hydrolysis, which is recognised a rate-limiting step in the AD process (Cirne *et al*., 2007), and also facilitate mathematical models limited through a lack of species differentiation in specific functional roles (Ramirez *et al*., 2009). Molecular fingerprinting methods have allowed important information to be gathered regarding community succession in a low-temperature context, primarily focused on methanogens. Recently however, a holistic polyphasic approach to bioprocess monitoring has begun to open the ecophysiological black box of engineered systems (Marzorati *et al*., 2008). With the essential aim to correlate key performance variables with community structure, novel microbial ecology techniques have been employed in a LTAD context, which include metagenomic (Schluter *et al*., 2008) and metaproteomic (Abram *et al*., 2011) approaches. Also, Werner *et al*. (2011) recently employed a deep sequencing time-series in nine full-scale AD plants and reported that the bacterial communities were very stable and unique to particular sites, which may represent a useful tool in LTAD studies. Therefore, the integration of well established community-profiling methodologies (qPCR, clone libraries) with function-based approaches (metaproteomics, RNA-derived DGGE, radiolablelled substrate uptake) may uncover the important functional groups within LTAD microbial communities, including the key metabolic processes been undertaken.

1.3 Molecular Microbial Ecology Techniques and Wastewater Treatment

1.3.1 16S rRNA gene analysis of community structure

It has been well documented that the majority of microbial species are ‘uncultured’ and do not grow under laboratory conditions (Handelsman, 2004; Alain & Querellou, 2009; Lewis *et al*., 2010). Thus, alternative methods have been developed to uncover microbial dynamics underpinning environmental samples. This progression in environmental microbiology has been facilitated by the development of the polymerase chain reaction (PCR), which was initially developed by Kary

Mullins in 1983. This approach revolutionised molecular-based microbiology, allowing culture-independent techniques to prosper in the identification and charcterisation of microorganisms isolated from environmental matrices.

Pioneering work by Woese (1987), revealed that organisms could be classified based on comparative sequence analysis of their small subunit ribosomal RNA, such as 16S rRNA. A phylogenetic framework, based on 16S rRNA sequences, was thus developed and could be used for specific primer and probe design and application (Kolbert & Persing, 1999; Palys *et al*., 1997). The reason behind the 16S rRNA genes success in PCR-based microbial ecology application is due to this gene possessing highly conserved regions, allowing not only bacterial communities to be analysed but archaeal communities also. The level of conservation is possibly due to the importance 16S rRNA in cellular function (Clarridge, 2004). Variability in the genetic sequence of 16S rRNA confers specificity to particular microbial groups, thus allowing comparative analysis to be undertaken via an online database of known 16S rRNA sequences, such as GenBank (publicly available nucleotide sequences for >250 000 formally described species; Benson *et al*., 2012).

Prior to database manipulation and interpretation, the amplified genes from environmental samples must be separated with molecular approaches applied. Such approaches include clone library generation, supported by amplified rRNA restriction analysis (ARDRA), and denaturing gradient gel electrophoresis (DGGE), which were employed during this study for the separation and characterisation of PCR products from bioreactor biomass samples. Also, real-time PCR was applied for the direct and quantitative amplification of variable regions of the 16S rRNA gene specific for targeted microbial groups.

1.3.1.1 Clone library analysis

Cloning and gene library construction have been applied to wastewater treatment systems in combination with other molecular microbial ecology techniques (Sanz & Köchling, 2007). The crux of genetic cloning involves the incorporation of

recombinant DNA into a host cell, with the subsequent growth, selection and screening of clones with desired DNA inserts and biological properties.

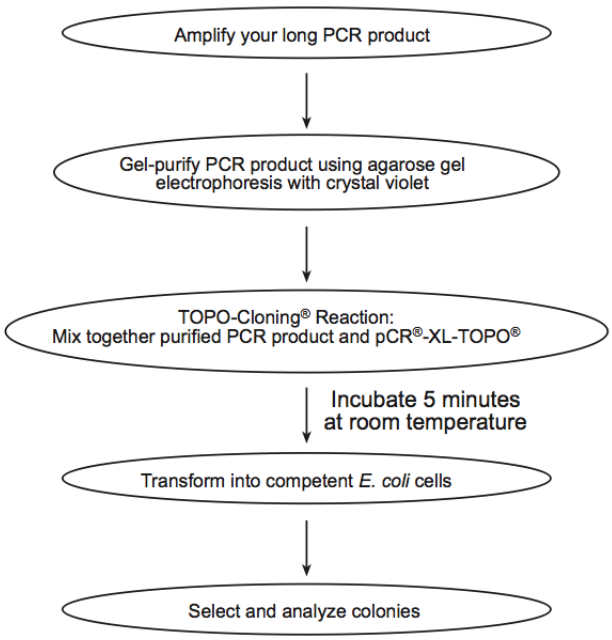


Figure 1.4 A flow chart outlining the experimental steps required to clone a long PCR product using a TOPO® XL PCR Cloning Kit (Invitrogen).

DNA Cloning of 16S rRNA gene fragments using host *Escherichia coli* cells presents one of the most widely used methods to separate PCR-amplified DNA, which are identical in length but varying in sequence (von Wintzingerode *et al*., 1997). Once individual PCR products have been cloned (Fig. 1.4), a transformed competent *E.coli* cell will contain a single 16S rDNA gene product from an organism originating from the environmental sample analysed. Screening of a 16S rRNA gene library can be undertaken through a number of methods, which include sequence-tagged site (STS)-PCR (Green & Olson, 1990), interspersed repetitive sequences (IRS)-PCR (Liu *et al*., 1995), amplified fragment length polymorphism (AFLP; Vos *et al*., 1995), and island rescue PCR (IRP; Valdes *et al*., 1994).

Amplified rRNA restriction analysis (ARDRA) was the selected screening method used in this study. ARDRA presents a cost-effective method to identify clones with similar sequences, therefore avoiding repetitive sequencing (Toomey, 2002).

Although no information about the types of microorganisms is given, this method provides unique genetic signatures, thus facilitating efficient comparisons of clones in the library. It involves the restriction of the DNA containing the 16S rRNA gene sequences with a tetrameric ('4-cutter') endonuclease enzyme, for example, *HaeIII*, which cuts the cloned sequence into smaller fragments at any GC-CC site. The resulting gel electrophoresis fragments provides a profile or fingerprint specific to each 16S rRNA gene sequence analysed. Thus, similar profiles are subsequently grouped into operational taxonomic units (OTUs; Moyer *et al*., 1994) with a representative clone from each OTU group sequenced for identification.

1.3.1.2 Denaturing Gradient Gel Electrophoresis (DGGE)

DGGE is an efficient separation technique of similar length DNA sequences (e.g. 16S rRNA-derived PCR products), which may vary as little as a single base pair modification. A major advantage of this method is its potential to visually profile and monitor changes undergoing in microbial communities, with high versatility, reliability and reproducibility (Muyzer & Smalla, 1998).

This molecular fingerprinting approach can generate patterns of genetic diversity from complex microbial ecosystems (Muyzer *et al*., 1993). DGGE has been applied extensively in microbial ecology, with recent studies including sulfidogenic community analysis in an anaerobic bioreactor (Dar *et al*., 2007), microbial communities evident at a deep-sea hydrothermal site (Postec *et al*., 2005) and oil contaminated soil community profiling (Hamamura *et al*., 2006).

DGGE is based on the principal of decreased electrophoretic mobility of partially melted double-stranded DNA molecules in polyacrylamide gels (e.g. PCR generated 16S rRNA DNA), which contain a gradient of DNA denaturants (Muyzer & Smalla, 1998). This allows for the separation of DNA fragments of the same length by sequence dependent manipulation of their melting domains, as sequence variation

within these domains causes the melting temperature to differ, thus DNA fragments with different sequences will stop migrating at different positions in the gel. Each of the resulting DNA bands theoretically represents a separate microbial population present in your original environmental sample. Also, this technique can be employed to monitor community changes over time as some bands may disappear or decrease in intensity, while others may emerge, therefore equating to shifts in microbial community structure.

Recently, the application of RNA-based DGGE profiling has been employed in microbial ecology studies, for example, a study by Corgie *et al*. (2006), found that by applying this approach, specific differences in rhizosphere associated bacterial species were gathered from the general community profiles (DNA-based) in coordination with the active species profiles (RNA-based). This implies that a linked DNA- and RNA- based molecular approach holds particular scope for unraveling the structural (DNA) and functional (RNA) diversity of microbial communities in the environment, particularly when operation parameters such as temperature are variable. Clear differences between DNA/RNA DGGE profiles have been recorded from environmental samples, such as from soil taken from Yellowstone national park (Norris *et al*., 2002; Fig. 1.5). A combined DNA- and RNA- based approach was was also employed in this study.

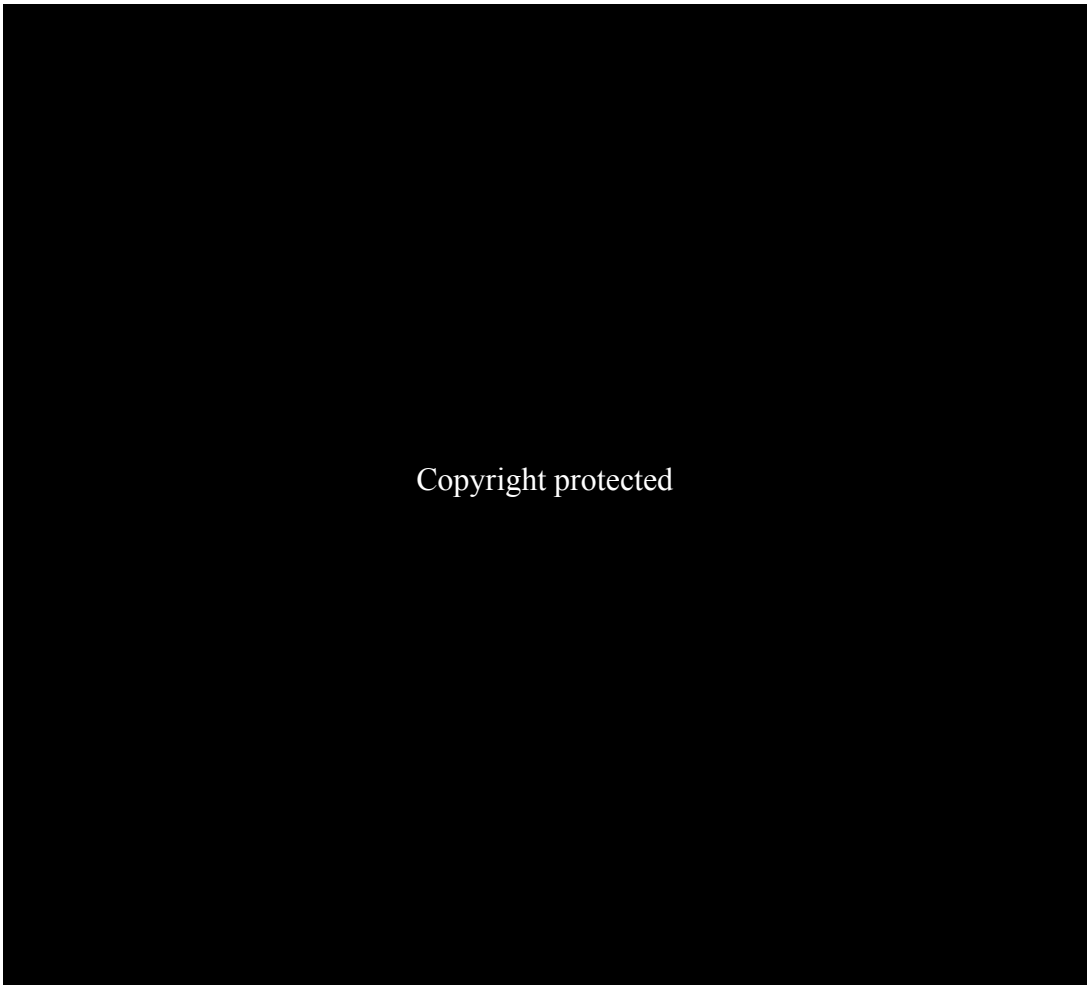


Figure 1.5 DGGE gel with (A) RNA-derived profiles and (B) DNA-derived profiles from soil samples taken from Yellowstone national park over time. Clear differences between profiles are present with a clear band evident on week three in RNA-derived profile (indicated by arrow) not present in DNA profile (Norris *et al*., 2002).

1.3.1.3 Phylogenetic analysis

Molecular techniques such as DGGE and clone library analysis can produce an array of DNA sequences, which require further coordination and analysis prior to data interpretation. These sequences can be compared to previous entries in a database, such as GenBank. The computational algorithm tool BLASTn (Altschul *et al*., 1997), can be used to search against an internal database of 16S rRNA sequences. The subsequent construction of multiple sequence alignments with software such as ClustalX (Thompson *et al*., 1997), allows for efficient organisation, visualisation, and analysis of this sequence data. This approach is required to compromise between recording a maximum amount of homology (matching sequences), while introducing

the smallest number of gaps by incorporating insertion/deletion events (Pace *et al*., 1986).

Thereafter, the closely related groups of sequences are aligned first with subsequent coordination of remaining sequence alignments in a phylogenetic tree, used to determine the position of the environmental sequence of interest. A phylogenetic tree is a binary tree representation of the resulting relationships between sequences used. Phylogenetic tree construction methods are widely accepted to fall into one of two categories: distance based and character based. These two categories both offer a variety of options when constructing trees in two different directions. The most common distance based methods are the unweighted pair group method using arithmetic averages (UPGMA; Sneath & Sokal, 1973) and Neighbor Joining (Saitou & Nei, 1987) algorithms, which are both based on the initial creation of a distance matrix. In theory, distance-based methods are the simplest way to construct a phylogenetic tree, with pairwise comparisons of a set of aligned sequences used to construct a distance matrix, which can then be converted into a bifurcating phylogenetic tree by grouping the most closely related pairs of sequences (Head *et al*., 1998). As such, a distance-based analyses was used for the construction of all phylogenetic trees in this study.

1.3.1.4 Quantitative PCR

Quantitative real-time polymerase chain reaction methodologies have been employed extensively in environmental microbiology, allowing for the successful quantification of microbial gene copy numbers (Becker *et al*., 2000; Lopez-Gutierrez *et al*., 2004; Zhang & Fang, 2006). In conjunction with primers designed to amplify the targeted gene sequences, this approach employs an additional fluorescent probe, which hybridises with a specific site within the targeted region. Each sequence of the primer and probe is designed to be specific for a target organism or microbial group, which equates to increased sensitivity and specificity with three oligonucleotides complementary to the target DNA sequence simultaneously employed (Yu *et al*., 2005). The amount of emitted fluorescence is directly proportional to the amount of

the targeted PCR product.

Two methods are usually employed, (i) SYBR green (Wittwer *et al*., 1997), and (ii) *Taq*Man (Holland *et al*., 1991). Each method requires an assay using a range of standards, which are constructed from known amounts of the target gene in question. The *Taq*Man assay (Applied Biosystems, Foster City, CA, USA) uses the 5' nuclease activity of *Taq* polymerase to cleave a non-extendible hybridisation probe during the extension phase of PCR (Heid *et al*., 1996). Dual labelled fluorogenic hybridisation probes (Bassler et al., 1995) are used, one acting as a reporter (e.g. FAM i.e. 6-carboxyfluoroscein) and the other acting as a quencher (e.g. TAMRA i.e. 6-carboxy-tetramethylrhodamin). After the probe is cleaved, the quencher dye no longer suppresses the reporter dye, thus the fluorescence can be detected. The Sybr Green assay (Molecular Probes, OR, USA) uses an intercalating SYBR Green I dye to bind to any double-stranded DNA and subsequently fluoresces as a result of this. Although this approach is cost effective, the level of false positive results is increased with this approach due to non-specific binding (Howell *et al*., 1999). Due to the sequence information of the organisms investigated in this study being known (Lee *et al*., 2008; Yu *et al*., 2005), with specific probe/primer sets already designed, the *Taq*Man assay was used in this study.

1.3.2 PCR-independent techniques

It is known that PCR-based methods may be affected by preferential or differential amplification (Reysenbach *et al*., 1992; Walsh *et al*., 1992), which may hinder the detection of some genotypes when analysing DNA extracted from an environmental matrix. Preferential PCR amplification may be caused by (i) primer mismatches at the annealing sites of the DNA templates of some genotypes or (ii) a lower rate of primer hybridisation to certain templates due to differential denaturation of these templates (Walsh *et al*., 1992). Systematically testing different sets of primers and enhancing DNA denaturation during PCR by using different reagents (denaturants and co-solvents) may solve these problems (Reysenbach *et al*., 1992). A second type of bias that may affect PCR carried out on complex bulk DNA extracts is the occurrence of (i) heteroduplexes that arise in later PCR cycles when primer

concentration decreases and the concentration of the PCR products are high enough to compete with the primers for annealing (Kanawaga, 2003), and (ii) chimeric amplicons that also form in later PCR cycles when the concentration of incompletely extended primers is high enough to compete with the original primer for annealing, or (iii) when template concentration is high enough to allow the re-annealing of templates before or during primer extension (Kanawaga, 2003). All these artifacts can generate additional signals that do not correspond to genotypes in the sample.

However, Suzuki *et al*. (1998), discussed that the possibility of bias due to reannealing in PCR may be small in environmental DNA samples, which are composed of highly diverse DNA templates, thus saturation point may not be reached for a single template. Indeed, there has been many suggestions put forward for minimising PCR artifacts in environmental profiling and fingerprinting approaches, which include using a small number of PCR amplification cycles (e.g. until a band is barely visible on agarose gels) to minimise chimeras and *Taq* DNA polymerase errors (Polz & Cavanaugh, 1998). Also, to minimise the presence of heteroduplex molecules, a reconditioning PCR step (e.g. three to five additional PCR cycles using fresh reagent mixture) may be included (Thompson *et al*., 2002).

Nevertheless, efficient PCR-independent methodologies are important, as they can facilitate the recovery of important cellular information as stand alone studies, or be integrated into a polyphasic approach with PCR-based approaches. Recent advances in molecular ecological techniques are evident, with the primary focus on the functional characterisation and quantification of gene products (mRNA, proteins and metabolites), as well as their interactions (Fig. 1.6). A number of such PCR-independent approaches are now reviewed.

Figure 1.6 Schematic diagram of various 'omics' technologies targeting different layers of cellular information (Zhang *et al*., 2010).

1.3.2.1 Physiological analysis

1.3.2.1.1 Specific methanogenic activity (SMA) assay

Methodologies for characterising the functions of microbial consortia present in environmental regimes are becoming increasingly important for monitoring purposes including, for example, in anaerobic digestion technology (Yu *et al*., 2005). Activity-based methods provide key insights into metabolic routes, such as the anaerobic process of methanogenesis, by ascertaining the metabolic ratio between the production of methane via the decarboxylation of acetate or the reduction of CO_2 (Wilkie & Colleran, 1987). In addition, monitoring of methane production activity utilising different substrates allows us to trace the activity of different trophic groups throughout the duration of an investigation. In this thesis, the method of Colleran *et al*. (1992), was employed for the measurement of specific methanogenic activity against both soluble and gaseous substrates.

1.3.2.1.2 Biodegradation & toxicity assays

Measuring the biodegradation capacity of a particular substance is important in order to assess the fate and behavior of these particular substances prior to application in lab- and pilot- scale bioreactor applications. During a biodegradation test of organic substances employing anaerobic microbial consotrtias, CH_4 and CO_2 are produced as final decomposition products. The amount of CH_4 and CO_2 produced per weight unit of test item is calculated. If the carbon content of the test item is known the percentage of biodegradation can thus be calculated as the percentage of solid carbon of the test item that has been converted to gaseous, mineral C.

However, the complexity of some substances being degraded can demonstrate inhibitory characteristics. As anaerobic digestion is undertaken by a complex mix of different microbial trophic groups, the inhibition of any trophic group involved in anaerobic digestion will impact the overall process. Specific investigations, targeting the effect of such a toxicant on individual steps of the overall process, are therefore more effective, and yield more information into the vulnerability of different microbial species. In practice, toxicity assays employ a range of SMA based batch assays, with replicate vials containing a range of toxicant concentrations, to determine the concentration that results in a 50% inhibition of a control methanogenic activity, using a range of substrates (Colleran *et al*., 1992).

Siggins *et al*. (2011c) demonstrated the successful application of a toxicity assay, measuring the capacity for anaerobic microbial communities to treat a wastewater supplemented with trichloroethylene (TCE) at low temperature (7°C). Here, the acetoclastic methanogen community showed an increased sensitivity to the presence of TCE and its degradation derivatives during the later stages of bioreactor operation.

1.3.2.2 Proteomics

Proteomics was originally defined as ‘‘the large-scale study of proteins expressed by an organism’’ (Wilkins *et al*., 1995), and has emerged as a promising technique used to characterise microbial activities at the molecular level. Since the 1990s, proteomics has become much more accessible in molecular research due to particular advances, which include: (i) the increase of genomic and metagenomic data, providing a solid basis for protein identification; and (ii) progression in sensitivity and accuracy of mass spectrometers, thus enabling high throughput protein identification. For example, isobaric tags for relative and absolute quantification (iTRAQ), allows up to eight samples to be compared through differentially labelled peptides. This proteomic method has been used in many studies, including investigations of bacteria grown under nitrate stress (Redding *et al*., 2006) and a methanogenic syntrophic co-culture (Walker *et al*., 2012). Moreover, improvements in computing power and bioinformatics have allowed substantial datasets to be processed and evaluated. Global analyses of protein expression has facilitated a holistic approach to characterise microbial metabolic dynamics in pure cultures and also from environmental samples (Fig. 1.7).

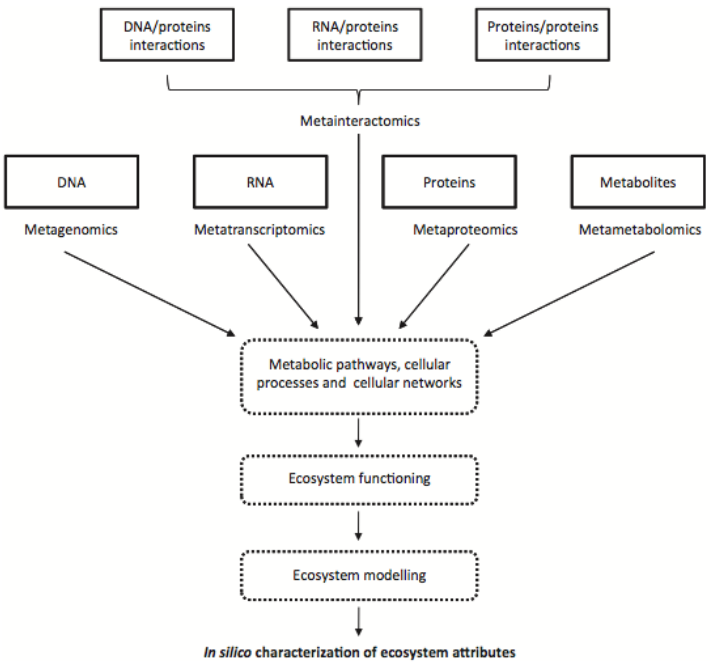


Figure 1.7 Schematic diagram of a system approach for the characterisation of microbial ecosystems (Siggins *et al*., 2011).

System-based approaches provide an emerging field of research in microbial ecology, whereby all levels of biological information are investigated (DNA, RNA, proteins and metabolites; Fig. 1.7) to capture the functional interactions occurring in a given ecosystem and identify characteristics that could not be accessed by the study of isolated components (Raes & Bork, 2008).

Metaproteomics, which is the identification of all the proteins expressed at a given time within an ecosystem (as defined by Wilmes & Bond, 2004), is an indispensable element of system approaches and plays a key role in the determination of microbial functionality. Microbial metaproteomics has been applied in the context of diverse environments such as soil (Benndorf *et al*., 2007; Williams *et al*., 2010), sediments (Benndorf *et al*., 2009; Bruneel *et al*., 2011), marine (Morris *et al*., 2010; Sowell *et al*., 2011), freshwater (Ng *et al*., 2010), human intestinal tract (Verberkmoes *et al*., 2009; Rooijers *et al*., 2011), human oral cavity (Rudney *et al*., 2010), animal guts (Toyoda *et al*., 2009) and natural and bioengineered systems (Ram *et al*., 2005; Wilmes *et al*., 2008; Jehmlich et al., 2010), with examples of successful metaproteomic experiments from environmental samples shown in Table 1.3.

Table 1.3 Overview of selected metaproteomic studies

Environment	Number of peptides /proteins identified	Method	Database used	Reference
Human gut	NA/2214 proteins	LC-MS/MS[a]	2 unmatched human gut metagenomes, several genomes from gut inhabitants and sevreal non-human gut genomes	Verberkmoes *et al*. (2009)
Soil	NA/122 proteins	2D-PAGE[b], MALDI TOF/TOF MS/MS[c]	Complete NCBInr, bacterial entries NCBInr and fungal entries NCBInr	Chourey *et al*. (2010)
Marine	6533 peptides/1042 proteins	LC-MS/MS	SAR11 clade and specific microorganisms from Sargasso Sea metagenome as well as genomes fromsequenced isolates	Sowell *et al*. (2009)
Freshwater	NA/1824 proteins	1D-PAGE[b], LC-MS/MS	Matched metagenomes	Lauro *et al*., (2011)
Activated sludge	NA/5029 proteins	LC-MS/MS	Three distinct unmatched activated sludge metagenomes	Wilmes *et al*. (2008)
Anaerobic digestion	NA/202 proteins	2D-PAGE, LC-ESI-MS/MS[d]	Bacterial entries of the NCBInr database	Jehmlich *et al*. (2010)

[a] Liquid chromatography-tandem mass spectrometry
[b] One/two-dimensional polyacrylamide gel electrophoresis
[c] matrix-assisted laser desorption ionisation-time of flight
[d] electrospray ionisation

Typically, proteomic analysis in a pure or mixed community context involves up to seven main steps (Fig. 1.8), namely sample collection, recovery of the targeted fraction, protein extraction, protein separation and/or fractionation, mass spectrometry analysis, databases searches and finally data interpretation, whereby the expressed proteins and pathways identified are used to access information about organism and/or system functioning (for detailed descriptions of the methodologies involved, see Wilmes & Bond, 2006 and Verberkmoes *et al*., 2009).

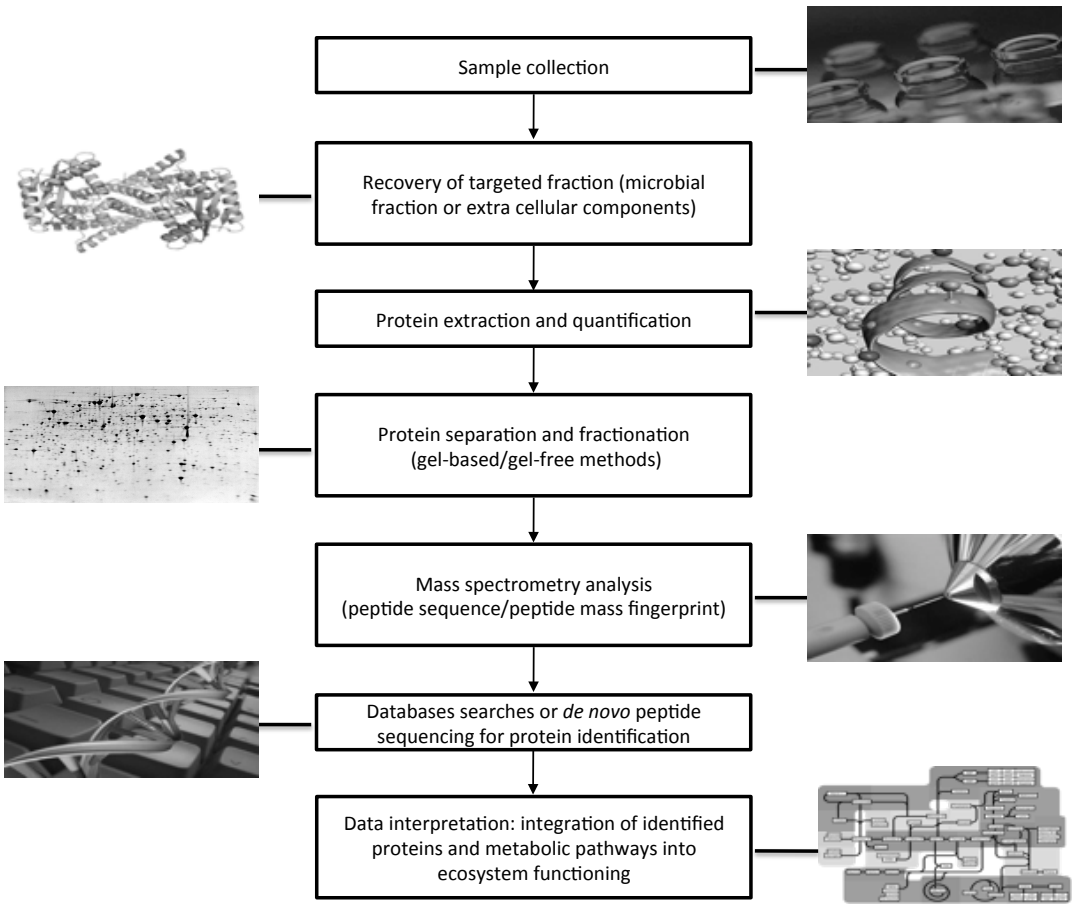


Figure 1.8 Typical workflow for proteomics analysis (Siggins *et al*., 2011).

Each microbial sample offers specific challenges and limitations within the proteomic workflow. In environmental samples, collection and recovery of the targeted fraction can be problematic, for example, with marine and freshwater, microbial samples can be recovered from hundreds of liters of water far away from laboratory facilities. The protein extraction step has proven specifically difficult when dealing with soil samples, which naturally contain interfering humic acids. Such compounds are usually co-extracted together with proteins and are known to interfere with protein quantification, separation and identification (Bastida *et al*., 2009). The use of gel-based methods for protein separation presents some disadvantages regardless of the origin of the sample. Such drawbacks are typically those associated with two-dimensional gel electrophoresis (2-DGE): proteins with extreme isoelectric points (basic or acidic) or extreme molecular weight (very large or very small), lipophilic proteins and low abundance proteins are typically excluded (Gygi *et al*., 2000; Ong & Pandey, 2001). A successful 2-DGE approach to anaerobic digestion granules was achieved by Abram *et al*. (2009).

Successful proteomics research relies on the availability of relevant genome or metagenome sequences when searching generated mass spectra against existing databases for protein identification. However, when no relevant sequences are available, *de novo* peptide sequencing can be used for protein identification (Lacerda *et al*., 2007; Fig. 1.8). In addition to being an integrative component of system approaches (Fig. 1.7), proteomics presents some valuable advantages for functional analyses. In this respect, proteomics was employed in this research to characterise community functioning underpinning anaerobic bioreactors (2D-LC-MS/MS) and also in a pure culture context to uncover low-temperature functional dynamics of a selected methanogenic microorganism (iTRAQ).

1.3.2.3 Microscopy techniques

Molecular techniques that do not rely on nucleic acid or protein extraction confer a structurally non-invasive approach to analyse microbial community structure, for example, in anaerobic biofilms. Microscopy methodologies such as *in situ* hybridisation offer a sensitive approach for quantitative analysis, as single cells can be specifically detected and counted under a microscope (Amann *et al*., 1995; Wintzingerode *et al*., 1997).

Prokaryotic cells can be identified without prior cultivation through the application of fluorescence *in situ* hybridisation (FISH), using rRNA targeted oligonucleotode probes (Wagner *et al*., 2003). This method can ultimately reveal the spatial relationships between microorganisms within a particular environment (Amann *et al*., 1997). The 3D structure or architecture of a community can be visualised by means of confocal laser scanning microscopy (CLSM), which involves the excitation of a fluorophore by a focused laser beam (Wilderer *et al*., 2002), which provides greater focal depth and the high resolution required to produce *in situ* imaging of microbial groups. FISH has been used extensively in anaerobic biofilm studies, for example, Collins *et al*. (2005), investigated *Crenarchaeota* in eight different anaerobic sludge samples, with close association with methanogenic species recorded.

The functional characterisation of environmental microbial groups, such as many methanogenic archaea in anaerobic biofilms is important, with methods such as FISH linked with direct functional approaches necessary to uncover discrete community dynamics. One complimentary method, which provides such useful functional information, is microautoradiography (MAR), which evaluates active uptake of a radiolabelled substrate by targeting phylogenetic groups identified through FISH analysis (Okabe *et al*., 2004). The combined MAR-FISH approach is an efficient technique, which enables quantifiable information relating to the ecophysiology of microorganisms at a single cell level in mixed communities to be recorded (Nielsen *et al*., 1999; Nielsen & Nielsen, 2005).

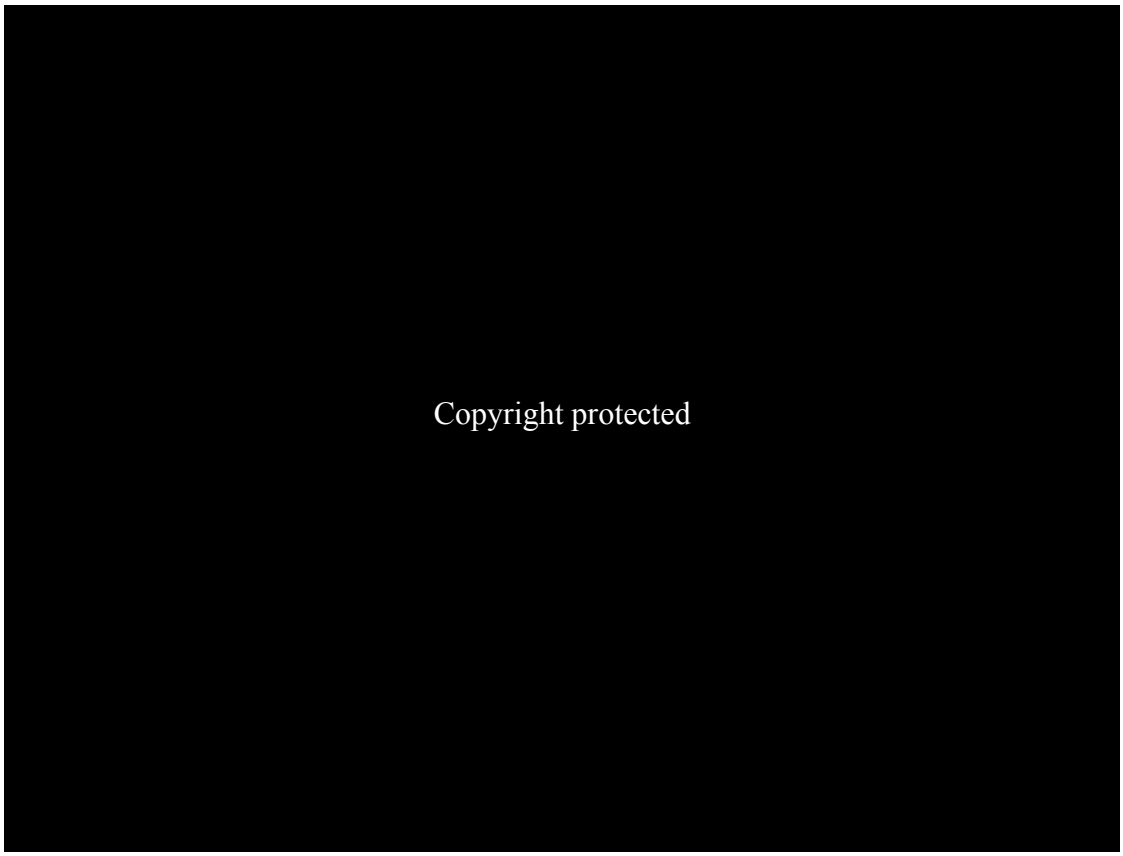


Figure 1.9 MAR-FISH workflow (Dumont & Murell, 2005)

The typical MAR-FISH procedure is composed of five key steps: (i) incubation with radiolabelled substrates, (ii) fixation and handling, (iii) FISH probe application, iv) autoradiographic development procedure (MAR); and finally, (v) microscopic observation (Fig. 1.9).

The radiotracers usually applied in MAR-FISH are typically the soft beta emitters ^{3}H (Ohta *et al*., 1996), ^{14}C (Lee *et al.*, 1999) and ^{33}P (Bücking & Heyser, 2001). Common to all radiotracers is the formation of precipitates in the form of silver grains on top and near the radiolabelled microorganism, which can be subsequently visualised through bright field or phase contrast microscopy (Nielsen & Nielsen, 2005).

MAR-FISH has been applied in a variety of studies analysing the ecophysiology of microorgansims in environmental settings, such as freshwater sediment (Gray *et al*.,

2000), activated sludge (Nielsen *et al*., 2000), a saltern pond (Rossello-Mora *et al*., 2003) and pelagic marine water (Alonso & Pernthaler, 2006). Another study by Ginige *et al*. (2004), found that MAR-FISH independently confirmed that a specific group of methanol oxidisers, DEN67, were the dominant bacterial group capable of anoxic [^{14}C]-methanol uptake in a methanol fed sequence batch reactor.

A particular focus for MAR-FISH studies has been the activated sludge process, with the ecophysiology of filamentous groups such as *Microthrix* and *Thiotrix* being investigated (Andreasen & Nielsen, 2000). Although there has been significant research undertaken in this field of wastewater microbial ecology, there is a lack of experimental MAR-FISH analysis regarding microbial aggregate dynamics within anaerobic wastewater treatment systems. Although there are limitations to this approach, which include substrate cross-feeding (Okabe et al., 2004), and partial digestion of radiolabelled substrate (Nielsen & Nielsen, 2005), MAR-FISH presents a viable approach to allow the simultaneous *in situ* identification of substrate uptake patterns relating to particular functional groups underpinning anaerobic granular sludge, particularly low-temperature samples. This PCR-independent microbial ecology approach was thus applied in coordination with proteomics and PCR-based methods in elucidating the microbial dynamics present in low-temperature anaerobic granules, with particular focus on methanogenic substrate uptake and functioning.

Other isotope labeling approaches include stable isotope probing (SIP) techniques, which utilises 13C-labelled growth substrates to link microbial growth to function by selective recovery of the 'heavy' ^{13}C-labelled DNA (Radajewski *et al*., 2003; Wagner, 2004). The DNA recovered can then be subsequently analysed by molecular methods (16S rRNA clone libraries) together with ^{12}C DNA in order to identify the metabolically active groups, which incorporated the ^{13}C DNA into their genome. In anaerobic digestion, this approach has been used extensively, for example in investigating the major phylogenetic groups underpinning a glucose-degrading anaerobic digester (Ito *et al*., 2012), as well as uncovering novel acetate utilising bacteria (Ito *et al*., 2011).

1.3.3 Statistical analysis in Microbial Ecology

DNA sequence data are being accumulated at a massive rate by high-throughput technologies, such as pyrosequencing (Edwards *et al*., 2006), single-cell genome sequencing (Zhang *et al*., 2006) and metagenomics (Venter *et al*., 2004). Efficient interpretation techniques are required to facilitate the evaluation of this information in the context of environmental parameters governing microbial community changes, which had been recorded from genetic fingerprinting methods (Hughes *et al*., 2001; Ramette, 2007).

Multivariate statistics have been demonstrated as being valuable with respect to environmental samples, for example, Rudi *et al*. (2004) used partial least square regression (PLSR) to relate physical/chemical properties to microbial community composition. Recently, multivariate statistics has been employed for the analysis of microbial communities originating from anaerobic bioreactors, including a study undertaken by O'Reilly *et al*. (2010), where DGGE bands were statistically analysed by non-metric multidimensional scaling (NMS). A concise review of multivariate analysis in microbial ecology is provided by Ramette (2007).

In this study, hierarchal cluster analysis was employed, which allows the visual grouping of multiple environmental samples. Unweighted Pair Group Method with Arithmetic mean (UPGMA; also known as group average) was the first strategy to account for group structure (Sokal & Michener, 1958), and it, along with Ward's method are considered to be the most appropriate clustering methods for the generation of dendrograms from fingerprint data (McCune & Grace, 2002). Correspondingly, the use of UPGMA cluster analysis to monitor microbial community dynamics in anaerobic bioreactors has been widely reported (Lee *et al*., 2008; Connaughton *et al*., 2006).

In addition to nucleic acid statistical enumeration, proteomic data interpretation is becoming more important in microbial ecology research, with distinct mathematical tools used to uncover the microbial functional profiles present in pure cultures and

environmental samples. Once peptides from tandem mass spectra have been recovered algorithms, such as the Paragon Algorithm (Shilov *et al*., 2007) are used to identify peptides in a sample, and through inference, determine which proteins have been detected (Fig. 1.10).

The Paragon Algorithm relies on three key innovations, (i) the likely relevance that each sequence segment of a database to the MS/MS spectrum is quantified on a continuum using many weighted *de novo* sequence tags to compute a Sequence Temperature Value (STV), (ii) feature probabilities are used to model the frequencies of peptide features such as modifications, digestion events, and substitutions and finally, (iii) an overall threshold is applied to the net effect of STV and feature probabilities, yielding a highly selective triage of which peptide hypotheses are worth scoring. This algorithm was used throughout the proteomic experimental analysis in this study.

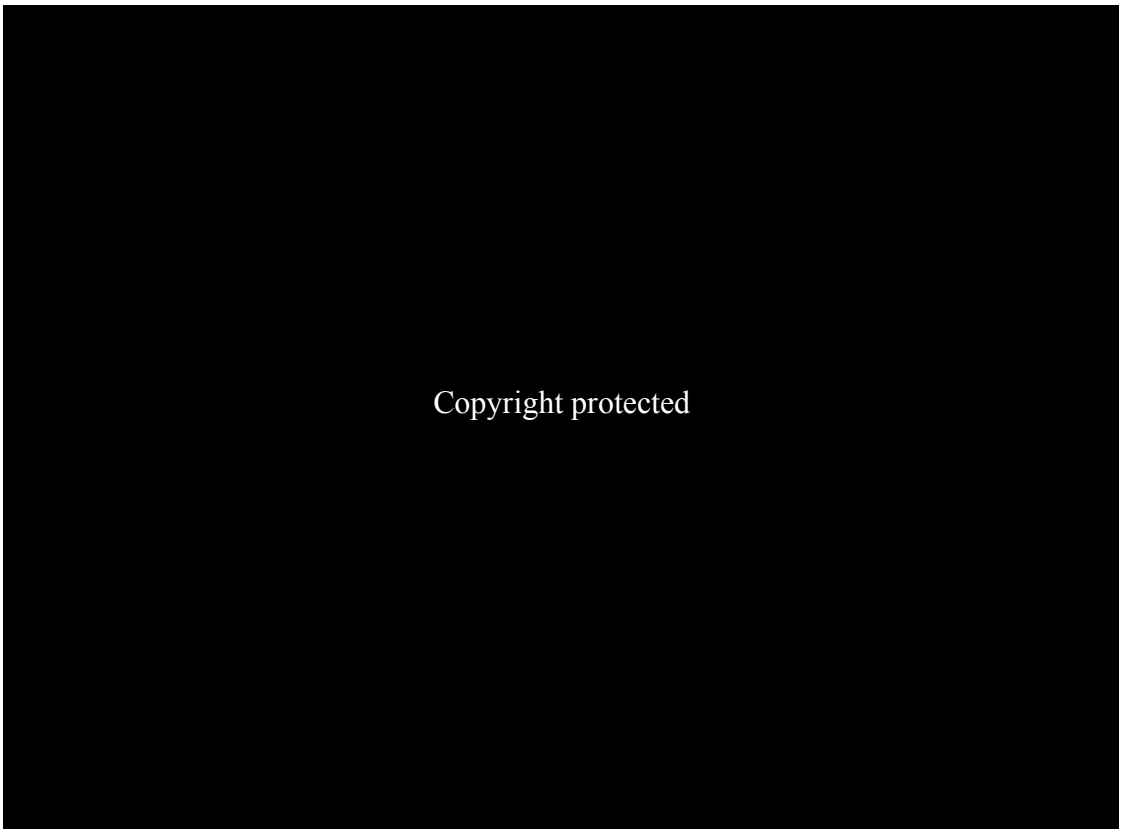


Figure 1.10 Paragon Algorithm workflow (Shilov *et al*., 2007)

1.4 Scope of this thesis

The main objective of this study was to investigate the key functional groups underpinning anaerobic bioreactors at low-temperature, with particular emphasis on methanogenic archaea.

In summary, 16S rRNA gene PCR-based approaches (clone libraries, DGGE, qPCR) were employed, in conjunction with metaproteomics to correlate important functional groups with community structure in relation to low-temperature anaerobic digestion. Physiological analysis was also applied via MAR-FISH, with bioreactor process information also linked in with molecular characterisation methods. Finally, a pure culture representative of a functional candidate present in low-temperature anaerobic digestion was investigated through proteomic analysis with key insights into low-temperature adaptation of methanogens functioning in anaerobic digestion recorded.

Chapter 2 investigates key microbial functional groups present in lab-scale anaerobic digestion bioreactors. Biomass samples were taken from three expanded granular sludge bed (EGSB; Fig. 1.3) bioreactors, R1 (37°C), R2 (15°C) and R3 (7°C), which were originally employed to treat a volatile fatty acid-based wastewater (organic loading rate 3kg COD m^{-3} day^{-1}). Metaproteomics in coordination with 16S rRNA gene phylogenetic approaches (clone libraries, qPCR) were used with changes in microbial community composition and metaproteomic profiles recorded as a function of temperature.

Chapter 3 characterises the sub-mesophilic functional characteristics of *Methanosarcina barkeri*, with a representative environmental homolog found to be functionally active in LTAD biomass. This was undertaken through a pure culture proteomic (iTRAQ, 2-DGE) approach with two experimental setups A) The first experimental setup aimed to characterise a low-temperature shock response of *M. barkeri* grown initially at 37°C with a temperature drop to 15°C, while B) the second experimental approach aimed to examine the low-temperature adaptation strategies

of this strain grown solely at 15°C, with Live/dead cell enumeration uncovered during sub-mesophilic growth.

Chapter 4 incorporates process information, molecular fingerprinting, ecophysiological charcterisation and metaproteomics to investigate methanogenic community structure and dynamics before, during and after a temperature drop to 15°C in replicate anaerobic bioreactors.

Chapter 5 presents the principle conclusions of this research, along with future recommendations.

1.5 References

Abram, F., Gunnigle, E. and O'Flaherty, V. (2009) Optimisation of protein extraction and 2-DE for metaproteomics of microbial communities from anaerobic wastewater treatment biofilms. *Electro*. **30**(23), 4149-4151.

Akila, G.and Chandra T.S. (2007) Performance of an UASB reactor treating synthetic wastewater at low-temperature using cold-adapted seed slurry Process. *Biochem*. **42**, 466-471.

Alain, K., Querellou, J. (2009) Cultivating the uncultured: limits, advances and future challenges. *Extremoph*. **13**(4), 583-594.

Alonso, C. and Pernthaler, J. (2006) Concentration-dependent patterns of leucine incorporation by coastal picoplankton. *Appl*. *Environ*. *Microbiol*. **72**, 2141-2147.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. **25**(17), 3389-3402.

Amann, R.I., Ludwig, W. and Schleifer, K.H. (1995) Phylogenetic identification and in-situ detection of individual microbial cells without cultivation. *Microbiol*. *Rev*. **59**, 143-169.

Amann, R., Glockner, F.O. and Neef, A. (1997). Modern methods in subsurface microbiology - in situ identification of microorganisms with nucleic acid probes. *FEMS Microbiol*. *Rev*. **20**, 191-200.

Andreasen, K. and Nielsen, P. H. (1997) Application of microautoradiography to the study of substrate uptake by filamentous microorganisms in activated sludge. *Appl*. *Environ*. *Microbiol*. **63**, 3662–3668.

Auman, A.J., Breezee, J.L., Gosink, J.J., Schumann, P., Barnes, C.R., Kämpfer, P. and Staley, J.T. (2010) *Psychromonas boydii* sp. nov., a gas-vacuolate, psychrophilic bacterium isolated from an Arctic sea-ice core. *Int*. *J*. *System*. *Evol*. *Microbiol*. **60**, 84-92.

Azam, F. and Malfatti, F. (2007) Microbial structuring of marine ecosystems. *Nat*. *Rev*. *Microbiol*. **5**, 782–791.

Bakermans, C., Ayala-del-Rio, H.L., Ponder, M.A., Vishnivetskaya, T., Gilichinsky, D., Thomashow, M.F. and Tiedje, J.M. (2006) *Psychrobacter cryohalolentis* sp. nov. and *Psychrobacter arcticus* sp. nov., isolated from Siberian permafrost. *Int*. *J*. *Syst*. *Evol*. *Microbiol*. **56**, 1285–1291.

Bassler, H.A., Flood, S.J.A., Livak, K.J., Marmaro, J., Knorr, R. and Batt, C.A. (1995) Use of a fluorogenic probe in a PCR-based assay for the detection of Listeria monocytogenes. *Am*. *Soc*. *Microbiol*. **61**(10), 3724-3728.

Bastida, F., Moreno, J.L., Nicolas, C., Hernandez, T. and Garcıa, C. (2009) Soil metaproteomics: a review of an emerging environmental science. Significance, methodology and perspectives. *J*. *Soil*. *Sci*. **60**, 845–859.

Becker, S., Böger, P., Oehlmann, R. and Ernst, A. (2000) PCR bias in ecological analysis: a case study for quantitative Taq nuclease assays in analysis of microbial communities. *Appl*. *Environ*. *Microbiol*. **66**(11), 4945-4953.

Beckmann, S., Leuders, T., Kruger, M., von Netzer, F., Engelen, B. and Heribert, C. (2011). *Appl*. *Environ .Microbiol*. **77**(11), 3749–3756.

Benndorf, D., Balcke, G.U., Harms, H. and Von Bergen, M. (2007) Functional metaproteome analysis of protein extracts from contaminated soil and groundwater. *ISME J*. **1**, 224–234.

Benndorf, D., Vogt, C., Jehmlich, N., Schmidt, Y., Thomas, H., Woffendin, G., Shevchenko, A., Richnow, H.H. and von Bergen, M. (2009) Improving protein extraction and separation methods for investigating the metaproteome of anaerobic benzene communities within sediments. *Biodeg*. **20**, 737–750.

Benson, D.A., Karsch-Mizrachi, I., Clark, K., Lipman, D.J., Ostell, J. and Sayers, E.O. (2012) Genbank. *Nucl*. *Acid*. *Res*. **40**(D1): D48–D53.

Bernardet, J.F., Segers, P., Vancanneyt, M., Berthe, F., Kersters, K. and Vandamme, P. (1996) Cutting a Gordian knot : emended classification and description of the genus *Flavobacterium*, emended description of the family *Flavobacteriaceae*, and proposal of *Flavobacterium hydatis* nom. nov. (basonym, *Cytophaga aquatilis* Strohl and Tait 1978). *Int*. *J*. *Syst*. *Bacteriol*. **46**, 128–148.

Bowman, J.P., McCammon, S.A., Lewis, T., Skerratt, J.H., Brown, J.L., Nichols, D. S. and McMeekin, T.A. (1998) *Psychroflexus torquis* gen. nov., sp. nov., a psychrophilic species from Antarctic sea ice, and reclassification of *Flavobacterium gondwanense* (Dobson et al. 1993) as *Psychroflexus gondwanense* gen. nov., comb. nov. *Microbiol*. **144**, 1601–1609.

Bruneel, O., Volant, A., Gallien, S., Chaumande, B., Casiot, C., Carapito, C., Bardil, A., Morin, G., Brown, G.E., Personne, C.J., le Paslier, D., Schaeffer, C., van Dorsselaer, A., Bertin, P.N., Elbaz-Poulichet, F. and Arsene-Ploetze, F. (2011) Characterization of the active bacterial community involved in natural attenuation processes in arsenic-rich creek sediments. *Microb*. *Ecol*. **61**, 793–810.

Bucking, H. and Heyser, W. (2001) Microautoradiographic localization of phosphate and carbohydrates in mycorrhizal roots of *Populus tremula* X *Populus alba* and implications for transfer pro- cesses in ectomycorrhizal associations. *Tree Physio*. **21**, 101– 107.

Cabeza, M.S., Baca, F.L., Puntes, E.M., Loto, F., Baigori, M.D. and Morata, V.I. (2011) Selection of psychrotolerant microorganisms producing cold-active pectinases for biotechnological processes at low-temperature. *Food Technol*. *Biotechnol*. **49**, 187-195

Cavicchioli, R., Thomas, T. and Curmi, P.M.G. (2000) Cold stress response in Archaea. *Extremoph*. **4**, 321-331.

Cavicchioli, R., Siddiqui, K.S., Andrews, D. and Sowers, K.R. (2002) Low-temperature extremophiles and their applications. *Curr*. *Opin*. *Biotechnol*. **13**, 253-261.

Cavicchioli, R., Ostrowski, M., Fegatella, F., Goodchild, A. and Guixa-Boixereu, N. (2003). Life under nutrient limitation in oligotrophic marine environments: an eco/physiological perspective of *Sphingopyxis alaskensis* (formerly *Sphingomonas alaskensis*). *Microb*. *Ecol*. **45**, 203–217.

Cavicchioli, R. (2006) Cold-adapted archaea. *Nat*. *Rev*. *Microbiol*. **4**, 331-343.

Cheng, S., Xing, D.F. and Logan, B.E. (2011) Electricity generation of single-chamber microbial fuel cells at low-temperatures. *Biosens*. *Bioelectron*. **26**(5), 1913-1917.

Chourey, K., Jansson, J., VerBerkmoes, N., Shah, M., Chavarria, K.L., Tom, E.L., Brodie, E.L. and Hettich, R.L. (2010) Direct cellular lysis/protein extraction protocol for soil metaproteomics. *J*. *Prot*. *Res*. **9**, 6615–6622.

Cirne, D.G., Paloumet, X., Bjornsson, L., Alves, M.M. and Mattiasson, B. (2006) Anaerobic digestion of lipid-rich waste—Effects of lipid concentration. *Renew*. *Ener*. **32**, 965–975

Clarridge, J.E. (2004) Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clin*. *Microbiol*. *Rev*. **17**(4), 840-862.

Colleran, E., Concannon, F., Golden, T., Geoghegan, F., Crumlish, B., Killilea, E., Henry, M. and Coates, J. (1992) Use of methanogenic activity tests to characterize anaerobic sludges, screen for anaerobic biodegradability and determine toxicity thresholds against individual anaerobic trophic groups and species. *Water Sci*. *Technol*. **25**(7), 31-40.

Collins, G., Foy, C., McHugh, S., Mahony, T. and O'Flaherty, V. (2005a) Anaerobic biological treatment of phenolic wastewater at 15-18°C. *Water Res*. **39**, 1614-1620.

Collins, G., O'Connor, L., Mahony, T., Gieseke, A., de Beer, D. and O'Flaherty, V. (2005b) Distribution, Localization, and Phylogeny of Abundant Populations of Crenarchaeota in Anaerobic Granular Sludge. *Appl*. *Environ*. *Microbiol*. **71**(11), 7523-7527.

Connaughton, S., Collins, G. and O'Flaherty, V. (2006) Psychrophilic and mesophilic anaerobic digestion of brewery effluent: a comparative study. *Water Res*. **40**, 2503 2510.

Corgie, S.C., Beguiristain, T. and Leyval, C. (2006) Profiling 16S bacterial DNA and RNA: Difference between community structure and transcriptional activity in phenanthrene polluted sand in the vicinity of plant roots. *Soil Biol*. *Biochem*. **38**, 1545–1553.

D'Amico, S., Collins, T., Marx, J.C., Feller, G. and Gerday, C. (2006) Psychrophilic microorganisms: challenges for life. *EMBO rep*. **7,** 385–389.

Dar, S.A., Yao, L., van Dongen, U., Kuenen, J.G. and Muyzer, G. (2007) Analysis of diversity and activity of sulfate reducing bacterial communities in sulfidogenic bioreactors using 16S rRNA and dsrB genes as molecular markers. *Appl*. *Environ*. *Microbiol*. **73**, 594–604.

Deming, J.W., Somers, L.K., Straube, W.L., Swartz, D.G. and MacDonald, M.T. (1988) Isolation of an barophilic bacterium and description of a new genus *Colwellia* gen. nov. *Syst.Appl.Microbiol*. **10**, 152-160.

Dumont, M.G. and Murrell, J.C. (2005) Stable isotope probing – linking microbial identity to function. *Nat*. *Rev*. *Microbiol*. **3**, 499–504.

Edwards, R.A., Rodriguez-Brito, B., Wegley, L., Haynes, M., Breitbart, M., Peterson, D.M., Saar, M.O., Alexander, S., Alexander, E.C. and Rohwer, F. (2006) Using pyrosequencing to shed light on deep mine microbial ecology. *BMC Genom*. **7**, 57.

Enright, A.M., Collins, G., and O'Flaherty, V. (2007) Low-temperature anaerobic biological treatment of toluene-containing wastewater. *Water Res*. **41**(7), 1465–1472.

Erkel, C., Kube, M., Reinhardt, R. and Liesack, W. (2006). Genome of Rice Cluster I archaea - the key methane producers in the rice rhizosphere. *Sci*. **313**, 370–372.

Feller, G. and Gerday, C. (2003) Psychrophilic enzymes: hot topics in cold adaptation. *Nat*. *Rev*. *Microbiol*. **1**, 200–208.

Feng, Y., Wang, X., Logan, B.E. and Lee, H. (2008) Brewery wastewater treatment using air–cathode microbial fuel cells. *Appl*. *Microbiol*. *Biotechnol*. **78**, 873–880.

Ferry, J.G. (1993) Methanogenesis: ecology, physiology, biochemistry & genetics. In Chapman & Hall microbiology series. New York.

Ferry, J.G. and Kastead, K.A. (2007) Methanogenesis. In: Cavicchioli R (ed) Archaea – molecular and cell biology. ASM Press, Washington, DC. 288-314.

Fey, A. and Conrad, R. (2000) Effect of temperature on carbon and electron flow and on the archaeal community in methanogenic rice field soil. *Appl*. *Environ*. *Microbiol*. **66**, 4790–4797.

Franzmann, P.D., Springer, N., Ludwig, W., Conway de Macario, E. and Rohde, M. (1992) A methanogenic archaeon from Ace Lake, Antarctica: *Methanococcoides burtonii* sp.nov. *Int*. *J*. *Syst*. *Appl*. *Microbiol*. **15**, 573-581.

Franzmann, P.D., Liu, Y., Balkwill, D.L., Aldrich, H.C., Conway de Macario, E. and Boone, D.R. (1997) *Methanogenium frigidum* sp. nov., a psychrophilic, H2-using methanogen from Ace Lake, Antarctica. *Int*. *J*. *Syst*. *Bacteriol*. **47**, 1068-1072.

Ginige, M.P., Hugenholtz, P., Daims, H., Wagner, M., Keller, J. and Blackall, L.L. (2004) Use of stable-isotope probing, full-cycle rRNA analysis, and fluorescence in situ hybridization–microautoradiography to study a methanol-fed denitrifying microbial community. *Appl*. *Environ*. *Microbiol*. **70**, 588–596.

Gosink, J.J., Herwig, R.P. and Staley, J.T. (1997) *Octadecabacter arcticus* gen. nov., sp. nov., and *O*. *antarcticus* sp. nov., nonpigmented, psychrophilic gas vacuolate bacteria from polar sea ice and water. *Syst*. *Appl*. *Microbiol*. **20**, 356–365.

Gosink, J.J., Woese, C.R. and Staley, J.T. (1998) *Polaribacter* gen. nov., with three new species, *P*. *irgensii* sp. nov., *P*. *franzmannii* sp. nov. and *P*. *filamentus* sp. nov., gas vacuolate polar marine bacteria of the Cytophaga-Flavobacterium-Bacteroides group and reclassification of '*Flectobacillus glomeratus*' as *Polaribacter glomeratus* comb. nov. *Int*.*J*.*Syst*.*Bacteriol*. **48**, 223-235.

Gray, N.D., Howarth, R., Pickup, R.W., Jones, J.G. and Head, I.M. (2000) Use of combined microautoradiography and fluorescence in situ hybridization to determine carbon metabolism in mixed natural communities of uncultured bacteria from the genus Achromatium. *Appl*. *Environ*. *Microbiol*. **66**, 4518–4522.

Green, E.D. and Olson, M.V. (1990) Systematic screening of yeast artificial-chromosome libraries by use of the polymerase chain reaction. *Proc*. *Natl*. *Acad*. *Sci*. **87**, 1213–1217.

Guo, H., Liu, R., Yu, Z., Zhang, H., Yun, J., Li, Y., Liu, X. and Pan, J. (2012) Pyrosequencing reveals the dominance of methylotrophic methanogenesis in a coal bed methane reservoir associated with Eastern Ordos Basin in China. *Int. J. Coal Geol.* **93**, 56–61.

Gygi, S.P., Rochon, Y., Franza, B.R. and Aebersold, R. (1999) Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* **19**, 1720–1730.

Hamamura, N., Olson, S.H., Ward, D.M. and Inskeep, W.P. (2006) Microbial population dynamics associated with crude oil biodegradation in diverse soils. *Appl. Environ. Microbiol.* **72**, 6316-6324.

Handelsman, J. (2004) Metagenomics: Application of genomics to uncultured microorganisms. *Microbiol. Mol. Biol. Rev.* **68**(4), 669-685

Head, I., Saunders, J.R. and Pickup, R.W. (1998) Microbial evolution, diversity and ecology: a decade of ribosomal RNA analysis of uncultivated organisms. *Microbiol. Ecol.* **35**, 1-21.

Heid, C.A., Stevens, J., Livak, K.J. and Williams, P.W (1996) Real time quantitative PCR *Genome Res.* **6**, 986-994.

Hirsch, P., Ludwig, W., Hethke, C., Sittig, M., Hoffmann, B. and Gallikowski, C.A. (1998) *Hymenobacter roseosalivarius* gen. nov., sp. nov. from continental Antarctica soils and sandstone: bacteria of the *Cytophaga*/*Flavobacterium*/*Bacteroides* line of phylogenetic descent. *Syst. Appl. Microbiol.* **21**, 374-383.

Hughes, J.B., Hellmann, J.J., Ricketts, T.H. and Bohannan, B.J.M. (2001) Counting the uncountable: Statistical approaches to estimating microbial diversity. *Appl. Environ. Microbiol.* **67**(10), 4399–4406.

Junge, K., Deming, J.W. and Eicken, H. (2004) Bacterial activity at -2 to -20°C in Arctic wintertime sea ice. *Appl. Environ. Microbiol.* **70**, 550-557.

Holland, P.M., Abramson, R.D., Watson, R. and Gelfand. D.H. (1991) Detection of specific polymerase chain reaction product by utilizing the 5'--3' exonuclease activity of Thermus aquaticus DNA polymerase. *Proc. Natl. Acad. Sci.* **88**, 7276-7280.

Howell, W.M., Jobs, M., Gyllensten, U. and Brookes, A.J. (1999) Dynamic allele-specific hybridisation. A new method for scoring single nucleotide polymorphisms. *Nat. Biotechnol.* **17**, 87-88.

Huang, L. and Logan, B.E. (2008) Electricity generation and treatment of paper recycling wastewater using a microbial fuel cell. *Appl. Microbiol. Biotechnol.* **80**, 349–355.

Inglis, G.J., Hayden, B.J. and Ross, A.H. (2000) An overview of factors affecting the carrying capacity of coastal embayments for mussel culture. NIWA Client Report CHC00/69, Christchurch, New Zealand.

Ito, T., Yoshiguchi, K., Ariesyady, H.D. and Okabe, S. (2011) Identification of a novel acetate-utilizing bacterium belonging to *Synergistes* group 4 in anaerobic digester sludge. *ISME*. *J*. **5**, 1844–1856.

Ito, T., Yoshiguchi, K., Ariesyady, H.D. and Okabe, S. (2011) Identification and quantification of key microbial trophic groups of methanogenic glucose degradation in an anaerobic digester sludge. *Bioresour*. *Technol*. **123**, 599-607.

Jehmlich, N., Kleinsteuber, S., Vogt, C., Benndorf, D., Harms, H., Schmidt, F., von Bergen, M. and Seifert, J. (2010) Phylogenetic and proteomic analysis of an anaerobic toluene-degrading community. *J*. *Appl*. *Microbiol*. **109**, 1937–1945.

Kanagawa, T. (2003). Bias and artifacts in multitemplate polymerase chain reactions (PCR). *J*. *Biosci*. *Bioeng*. **96**, 317-323.

Kolbert, C.P. and Persing, D.H. (1999) Ribosomal DNA sequencing as a tool foridentification of bacterial pathogens. *Curr*. *Opin*. *Microbiol*. **2**, 299-305.

Kotsyurbenko, O.R., Nozhevnikova, A.N., Soloviova, T.I. and Zavarzin, G.A. (1996) Methanogenesis at low-temperatures by microflora of tundra wetland soil. *Antonie van Leeuwenhoek*. **69**, 75–86.

Kotsyurbenko, O.R., Glagolev, M.V., Nozhevnikova, A.N. and Conrad, R. (2001) Competition between homoacetogenic bacteria and methanogenic archaea for hydrogen at low-temperature. *FEMS Microbiol*. *Ecol*. **38**, 153–159.

Kuddus, M. and Ramteke, P.W. (2009) Cold-active extracellular alkaline protease from an alkaliphilic *Stenotrophomonas maltophilia*: Production of enzyme and its industrial applications. *Can*. *J*. *Microbiol*. **55**, 1294-1301.

Lacerda, C.M., Choe, L.H. and Reardon, K.F. (2007) Metaproteomic analysis of a bacterial community response to cadmium exposure. *J*. *Prot*. *Res*. **6**, 1145-1152.

Lauro, F.M., DeMaere, M.Z., Yau, S., Brown, M.V., Ng, C., Wilkins, D., Raftery, M.J., Gibson, J.A., Andrews-Pfannkoch, C., Lewis, M., Hoffman, J.M., Thomas, T. and Cavicchioli, R. (2011) An integrative study of a meromictic lake Ecosystem in Antarctica. *ISME J*. **5**, 879–895.

Lay, J.J., Li, Y.Y. and Noike, T. (1998) Interactions between homoacetogens and methanogens in lake sediments. *J*. *Fermen*. *Bioeng*. **86**, 467–471.

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. *Appl*. *Environ*. *Microbiol*. **65**, 1289–1297.

Lee, C., Kim, J., Shing, S.G. and Hwang, S. (2008) Monitoring bacterial and archaeal community shifts in a mesophilic anaerobic batch reactor treating a high-strength organic wastewater. *FEMS Mcrobiol*. *Ecol*. **65**, 544-554.

Lettinga, G., Rebac, S. and Zeeman, G. (2001) Challenge of psychrophilic anaerobic wastewater treatment. *Trends Biotechnol*. **19**(9), 363-370.

Lewis, K., Epstein, S., D'Onofrio, A. and Ling, L.L. (2010). Uncultured microorganisms as a source of secondary metabolites. *J*. *Antibiot*. 1–9.

Lim, J., Thomas, T. and Cavicchioli, R. (2000) Low-temperature regulated DEAD-box RNA helicase from the Antarctic archaeon, *Methanococcoides burtonii*. *J*. *Mol*. *Biol*. **297**, 553–567.

Lin, C.Y., Noike, T., Sato, K. and Matsumoto, J. (1987) Temperature characteristics of the methanogenesis process in anaerobic digestion. *Water Sci*. *Technol*. **19**, 299–310.

Liu, J., Stanton, V.P., Fujiwara, T.M., Wang, J.X., Rezonzew, R., Crumley, M.J., Morgan, K., Gros, P., Housman, D. and Schurr, E. (1995) Large- scale cloning of human chromosome 2-specific yeast artificial chromosomes (YACs) using an interspersed repetitive sequences (IRS)-PCR approach. *Genom*. **26**, 178-191.

Liu, H., Ramnarayanan, R. and Logan, B. E. (2004) Production of electricity during wastewater treatment using a single chamber microbial fuel cell. *Environ*. *Sci*. *Technol*. **38**, 2281 – 2285.

Logan, B.E., Cheng, S.A., Watson, V. and Estadt, G. (2007) Graphite fiber brush anodesfor increased power production in air-cathode microbial fuel cells. *Environ*. *Sci*. *Technol*. **41**, 3341-3346.

Lopez-Gutierrez, J.C., Henry, S., Hallet, S., Martin-Laurent, F., Catroux, G. and Philppot, L. (2004) Quantification of a novel group of nitrate-reducing bacteria in the environment by real-time PCR. *J*. *Microbiol*. *Meth*. **57**(3), 399-407.

Madden, P., Chinalia, F.A., Enright, A.M., Collins, G. and O'Flaherty, V., (2010) Perturbation-independent community development in low-temperature anaerobic biological wastewater treatment bioreactors. *Biotechnol*. *Bioeng*. **105** (1), 79-87.

Martinez-Sosa, D., Helmreich, B., Netter, T., Paris, S., Bischof, F. and Harald, H. (2011) Anaerobic submerged membrane bioreactor (AnSMBR) for municipal wastewater treatment under mesophilic and psychrophilic temperature conditions. *Biores*. *Technol*. **22**, 10377-10385.

Marzorati, M., Wittebolle, L., Boon, N., Daffonchio, D. and Verstraete, W. (2008) How to get more out of molecular fingerprints: practical tools for microbial ecology. *Environ*. *Microbiol*. **10**, 1571–1581.

McCarty, P.L. and Smith, D.P. (1986) Anaerobic waste-water treatment 4. *Environ*. *Sci*. *Technol*. **20**(12), 1200–1206.

McCollom, T.M. (1999) Methanogenesis as a potential source of chemical energy for primary biomass pro- duction by autotrophic organisms in hydrothermal sys- tems on Europa. *J*. *Geophys*. *Res*. **104**, 30729–30742.

McCune, B. and Grace, J.B. (2002) Analysis of ecological communities. MJM Software Design, Gleneden Beach, Oregon.

McKeown, R., Scully, C., Enright, A.M., Chinalia, F.A., Lee, C., Mahony, T., Collins, G. and O'Flaherty, V. (2009) Psychrophilic methanogenic community development during long-term cultivation of anaerobic granular biofilms. *ISME*. *J*. **3**, 1231-1242.

McKeown, R., Hughes, D., Collins, G., Mahony, T. and O'Flaherty, V. (2012) Low-temperature anaerobic digestion for wastewater treatment. *Biotechnol*. **23**, 444-451.

Müller, V., Blaut, M. & Gottschalk, G. (1986). Utilization of methanol plus hydrogen by *Methanosarcina barkeri* for methanogenesis and growth. *Appl*. *Environ*. *Microbiol*. **52**, 269–274.

Moon, H., Chang, I.S. and Kim, B.H. (2006) Continuous electricity production from artificial wastewater using a mediator-less microbial fuel cell. *Bioresource Technol*. **97**, 621–627.

Morita, R.Y. (1975) Psychrophilic Bacteria. *Bacteriol Rev*. **39**, 144-167.

Moyer, C.L., Dobbs, F.C. and Karl, D.M. (1994) Estimation of diversity and community structure through restriction fragment length polymorphism distribution analysis of bacterial 16S rRNA genes from a microbial mat at an active hydrothermal vent system. Loihi, Seamount, Hawaii. *Appl*. *Environ*. *Microbiol*. **60**, 871-879.

Methe, B.A., Nelson, K.E., Deming, J.W., Momen, B., Melamud, E., Zhang, X.J.,Moult, J., Madupu, R., Nelson, W.C., Dodson, R.J., Brinkac, L.M., Daugherty, S.C., Durkin, AS., DeBoy, R.T., Kolonay, J.F., Sullivan, S.A., Zhou, L.W., Davidsen, T.M., Wu, M., Huston, A.L., Lewis, M., Weaver, B., Weidmann, J.F., Khouri, H., Utterback, T.R., Feldblyum, T.V. and Fraser, C.M. (2005) The psychrophilic lifestyle as revealed by the genome sequence of Colwellia psychrerythraea 34H through genomic and proteomic analyses. *Proc Natl Acad Sci USA*. **102**, 10913–10918.

Morris, R.M., Nunn, B.L., Frazar, C., Goodlett, D.R., Ting, Y.S. and Rocap, G. (2010) Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *ISME J*. **4**, 673–685.

Muyzer, G., deWall, E. and Uitterlinden, A.G. (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl*. *Environ*. *Microbiol*. **59**, 695-700.

Muyzer, G. and Smalla, K. (1998) Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. *Antonie van Leeuwenhoek*. **73**, 127-141.

Ng, C., DeMaere, M.Z., Williams, T.J., Lauro, F.M., Raftery, M., Gibson, J.A., Andrews-Pfannkoch, C., Lewis, M., Hoffman, J.M., Thomas, T. and Cavicchioli, R. (2010) Metaproteogenomic analysis of a dominant green sulphur bacterium from Ace Lake, Antartica. *ISME J*. **4**, 1002–1019.

Nielsen, A.T., Liu, W.T., Filipe, C., Grady, L., Molin, S., and Stahl, D.A. (1999) Identification of a novel group of bacteria in sludge from a deteriorated biological phosphorus removal reactor. *Appl*. *Environ*. *Microbiol*. **65**, 1251–1258.

Nielsen, P.H., Muro, M.A. and Nielsen, J.L. (2000) Studies on the in situ physiology of *Thiothrix* spp. present in activated sludge. *Environ*. *Microbiol*. **2**, 389–398.

Nielsen, P.H. and Nielsen, J.L. (2005) Microautoradiography: recent advances within the studies of the ecophysiology of bacteria in biofilms. *Water Sci*. *Technol*. **53**(7), 187-194.

Nogi, Y., Masui, N. and Kato, C. (1998) *Photobacterium profundum* sp. nov., a new, moderately barophilic bacterial species isolated from a deep-sea sediment. *Extremoph*. **2**, 1–7.

Norris, T.B., Wraith, J., Castenholz, R.C. and McDermott. T.R.(2002). Soil microbial community diversity after a recent geo- thermal heating event. *Appl*. *Environ*. *Microbiol*. **68**, 6300-6309.

Ohta, H., Nishikawa, H., Hirai, K., Kato, K. and Miyamoto, M. (1996) Relationship of impaired brain glucose metabolism to learning deficit in senescence-accelerated mouse. *Neurosci*. *Let*. **217**, 37-40.

Okabe, S., Kindaichi, T. and Tsukasa, I. (2004) MAR–FISH: an ecophysiological approach to link phylogenetic affiliation and in situ metabolic activity of microorganisms at a single-cell resolution. *Microb*. *Environ*. **19**, 83–98.

Ong, S.E. and Pandey, A. (2001) An evaluation of the use of two- dimensional gel electrophoresis in proteomics. *Biomol*. *Eng*. **18**, 195–205.

O'Reilly, J., Lee, C., Collins, G., Chinalia, F.A., Mahony. T. and O'Flaherty, V. (2009) Quantitative and qualitative analysis of methanogenic communities in mesophilically and psychrophilically cultivated anaerobic granular biofilms. *Water Res*. **43**, 3365-3374.

O'Reilly, J., Lee, C., Chinalia, F., Collins, G., Mahony, T. and O'Flaherty, V. (2010) Microbial community dynamics associated with biomass granulation in low-temperature (15°C) anaerobic wastewater treatment bioreactors. *Bioresour*. *Technol*. **101**, 6336-6344.

Pace, N.R., Stahl, D.A., Lane, D.J. and Olsen, G.J. (1986) The analysis of natural microbial populations by ribosomal RNA sequences. *Adv*. *Microbiol*. *Ecol*. **9**, 1-55.

Palys, T., Nakamura, L.K. and Cohan, F.H. (1997) Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. *Int*. *J*. *Sust*. *Bacteriol*. **47**, 1145-1156.

Polz, M. F. and Cavanaugh, C. M. (1998). Bias in template-to-product ratios in multitemplate PCR. *Appl Environ*. *Microbiol*. **64**, 3724–3730.

Postec, A., Urios, L., Lesongeur, F., Ollivier, B., Querellou, J. and Godfroy, A. (2005) Continuous enrichment culture and molecular monitoring to investigate the microbial diversity of thermophiles inhabiting deep-sea hydrothermal eco- systems. *Curr*. *Microbiol*. **50**, 138–144.

Pozo del, R., Diez, V., Garrido, S.E., Morales, M. and Osorio, R. (2002) Hydraulic distribution effect on a real-scale trickling filter. *Environ*. *Eng*. *Sci*. **19**(3), 151–157.

Preston, C.M., Wu, K.Y., Molinski, T.F. and DeLong, E.F. (1996) A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. *Proc*. *Natl*. *Acad*. *Sci*. *USA*. **93**, 6241-6246.

Purdy, K.J., Nedwell, D.B. and Embley, T.M. (2003) Analysis of the sulfate-reducing bacterial and methanogenic archaeal populations in contrasting Antarctic sediments. *Appl*. *Environ*. *Microbiol*. **69**, 3181–3191.

Raes, J. and Bork, P. (2008) Molecular eco-systems biology: towards an understanding of community function. *Nat*. *Rev*.*Microbiol*. **6**, 693–699.

Radajewski, S., McDonald, I.R. and Murrell, J.C. (2003) Stable-isotope probing of nucleic acids: a window to the function of uncultured microorganisms. *Curr*. *Opin*. *Biotechnol*. **14**, 296-302.

Ram, R.J., VerBerkmoes, N.C., Thelen, M.P., Tyson, G.W., Baker, B.J., Blake, R.C., Shah, M., Hettich, R.L. and Banfield, J.F. (2005) Community proteomics of a natural microbial biofilm. *Sci*. **308**, 1915–1920.

Ramirez, I., Volcke, E.I., Rajinikanth, R. and Steyer, J.P. (2009) Modeling microbial diversity in anaerobic digestion through an extended ADM1 model. *Water Res*. **43**(11), 2787-800.

Ramette, A. (2007) Multivariate analyses in microbial ecology. *FEMS Microbiol*. *Ecol*. **62**, 142–160.

Redding, A.M., Mukhopadhyay, A., Joyner, A.C., Hazen, T.C. and Keasling, J.D. (2006) Study of nitrate stress in *Desulfovibrio vulgaris* Hildenborough using iTRAQ proteomics. *Brief*. *Funct*. *Genom*. *Proteom*. **6**(15), 4321-4334.

Reysenbach, A.L., Giver, L.J., Wickham, G.S. and Pace, N.R. (1992) Differential Amplification of rRNA Genes by Polymerase Chain Reaction. *Appl*. *Environ*. *Microbiol*. **58**, 3417-3418.

Rogers, S. and McClure, N. (2003). The role of microbiological studies in bioremediation process optimisation. in: Singleton, I., Milner, M., and Head, I. (eds), Bioremediation: A Critical Review, Norfolk :Horizon Press: 27-60.

Rodrigues, D. F., Goris, J., Vishnivetskaya, T., Gilichinsky, D., Thomashow, M. F. & Tiedje, J. M. (2006). Characterization of Exiguobacterium isolates from the Siberian permafrost. Description of Exiguobacterium sibiricum sp. nov. *Extremoph*. **10**, 285–294.

Rooijers, K., Kolmeder, C., Juste, C., Dore, J., de Been, M., Boeren, S., Galan, P., Beauvallet, C., de Vos, W. and Schaap, P. (2011) An iterative workflow for mining the human intestinal metaproteome. *BMC Genom*. **12**, 6.

Rossello-Mora, R., Lee, N., Anton, J. and Wagner, M. (2003) Substrate uptake in extremely halophilic microbial communities revealed by microautoradiography and fluorescence in situ hybridization. *Extremoph*. **7**, 409-413.

Rudi, K., Maugesten, T., Hannevik, S.E. and Nissen, H. (2004) Explorative multivariate analyses of 16S rRNA gene data from microbial communities in modified-atmosphere-packed salmon and coalfish. *Appl*. *Environ*. *Microbiol*. **70**, 5010–5018.

Rudney, J.D., Xie, H., Rhodus, N.L., Ondrey, F.G. and Griffin, T.J. (2010) A metaproteomic analysis of the human salivary microbiota by three-dimensional peptide fractionation and tandem mass spectrometry. *Mol*. *Oral*. *Microbiol*. **25**, 38–49.

Russell, N.J., Harrisson, P., Johnston, I.A., Jaenicke, R., Zuber, M., Franks, F. and Wynn-Williams, D. (1990) Cold adaptation of microorganisms [and discussion]. *Phil*. *Trans*. *Royal Soc*. *London*. *Series B, Biological Sciences*. **336**, 595-611.

Sanz, J.L. and Köchling, T. (2007) Molecular biology techniques used in wastewater treatment: an overview. *Process*. *Biochem*. **42**, 119–133.

Saitou, N. and Nei, M. (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol*. *Biol*. *Evol*. **4**, 406–425.

Saravanane, R. and Murthy, D.V.S. (1999) Application of anaerobic fluidized bed reactors in wastewater treatment: a review. *Environ*. *Manag*. *and Heal*. **11**(2), 97-117.

Sattley, M.W. and Madigan, M.T. (2006) Isolation, characterization, and ecology of cold-Active, chemolithotrophic, sulfur-oxidizing bacteria from perennially ice-covered Lake Fryxell, Antarctica. *Appl*. *Environ*. *Microbiol*. 72(8), 5562–5568.

Shilov, I.V., Seymour, S.L., Patel, A.A., Loboda, A., Tang, W.H., Keating, S.P., Hunter, H.L., Nuwaysir ,L.M. and Schaeffer D.A. (2007) The Paragon algorithm, a next generation search engine that uses sequence temperature values andfeature probabilities to identify peptides from tandem mass spectra. *Mol*. *Cell*. *Proteom*. **6**, 1638–1655.

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

Siggins, A., Enright, A.M., O'Flaherty, V., 2011. Temperature dependent (37°C-15°C) anaerobic digestion of a trichloroethylen-contaminated wastewater. *Bioresource*. *Technol*. **102**, 7645-7656.

Siggins, A., Gunnigle, E. and Abram, F. (2012) Exploring mixed microbial community functioning: recent advances in metaproteomics. *FEMS Microbiol*. *Ecol*. **80**, 265–280.

Simankova, M.V., Kotsyurbenko, O.R., Lueders, T., Nozhevnikova, A.N., Wagner, B., Conrad, R. and Friedrich, M.W. (2003) Isolation and characterization of new strains of methanogens from cold terrestrial habitats. *Syst Appl Microbiol*. **2**, 312–318.

Sneath, P.H.A. and Sokal, R.R. (1973). Numerical taxonomy — the principles and practice of numerical classification. (W. H. Freeman: San Francisco.).

Sokal, R.R. and Michener, C.D. (1958) A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin*. **38**, 1409-1438.

Sotemann, S.W., van Rensburg, P., Ristow, N.W., Wentzel, M.C., Loewenthal, R.E. and Ekama, G.A. (2004) Integrated chemical/physical and biological processes modelling Part 2: Anaerobic digestion of sewage sludges. *Wat*. *Inst*. *South Afr*. *Conf*. *Proc*. 1327-1344.

Sowell, S.M., Wilhelm, L.J., Norbeck, A.D., Lipton, M.S., Nicora, C.D., Barofsky, D.F., Carlson, C.A., Smith, R.D. and Giovanonni, S.J. (2009) Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J*. **3**, 93–105.

Sowell, S.M., Abraham, P.E., Shah, M., Verberkmoes, N.C., Smith, D.P., Barofsky, D.F. and Giovannoni, S.J. (2011) Environmental proteomics of microbial plankton in a highly productive coastal upwelling system. *ISME J*. **5**, 856–865.

Stallwood, B., Shears, J., Williams, P.A. and Hughes, K. A. (2005) Low-temperature bioremediation of oil-contaminated soil using biostimulation and bioaugmentation with a *Pseudomonas* sp. from maritime Antarctica. *J*. *Appl*. *Microbiol*. **99**, 794–802.

Suzuki, M.T., Rappe, M.S. and Giovannoni, S.J. (1998) Kinetic bias estimates of coastal picoplankton community structure obtained by measurements of small-subunit rRNA gene PCR amplicon length heterogeneity. *Appl*. *Environ*. *Microbiol*. **64**, 4522 – 4529.

Tarpgaard, I.H., Boetius, A. and Finster, K. (2006) *Desulfobacter psychrotolerans* sp. nov., a new psychrotolerant sulfate-reducing bacterium and description of its physiological response to temperature changes. *Antonie van Leeuwenhoek*. **89**,109–124.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*. **25**(24), 4876-4882.

Thompson, J.R., Marcelino, L.A. and Polz, M.F. (2002) Heteroduplexes in mixed-template amplifications: formation, consequence and elimination by 'reconditioning PCR'. *Nuc. Acid. Res.* **30**, 2083 – 2088.

Toomey, M. (2002) Genetic analysis of anaerobic digestion biomass in the presence and absence of sulphate and nitrate. PhD Thesis, National University of Ireland, Galway.

Torsvik, V., Øvreås, L. and Thingstad, T.F. (2002) Prokaryotic diversity—Magnitude, dynamics, and controlling factors. *Sci.* **296**, 1064–1066.

Toyoda, A., Iio, W., Mitsumori, M. and Minato, H. (2009) Isolation and identification of cellulose-binding proteins from sheep rumen contents. *Appl. Environ. Microbiol.* **75**, 1667–1673.

van Trappen, S., Vandecandelaere, I., Mergaert, J. and Swings, J. (2004) *Gillisia limnaea* gen. nov., sp. nov., a new member of the family *Flavobacteriaceae* isolated from a microbial mat in Lake Fryxell, Antarctica. *Int. J. Syst. Evol. Microbiol.* **54**, 445- 448.

Valdes, J.M., Tagle, D.A. and Collins, F.S. (1994) Island rescue PCR: a rapid and efficient method for isolating transcribed sequences from yeast artificial chromosomes and cosmids. *Proc. Natl. Acad. Sci. USA.* **91**: 5377- 5381.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., Wu, D., Paulsen, I., Nelson, K.E. and Nelson, W. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Sci.* **304**, 66- 74.

Verberkmoes, N.C., Russell, A.L., Shah, M., Godzik, A., Rosenquist, M., Halfvarson, J., Lefsrud, M.G., Apajalahti, J., Tysk, C., Hettich, R.L. and Jansson, J.K. (2009) Shotgun metaproteomics of the human distal gut microbiota. *ISME J.* **3**, 179–189.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* **11**, 4407–4414.

Wagner, M., Horn, M. and Daims, H. (2003) Fluorescence *in situ* hybridisation for the identification and characterisation of prokaryotes. *Curr. Opin. Microbiol.* **6**, 302–309.

Wagner, M. (2004) Deciphering the function of uncultured microorganisms. *ASM News. 70*, 63-70.

Walker, C.B., Redding-Johanson, A.M., Baidoo, E.E., Rajeev, L., He, Z., Hendrickson, E.L., Joachimiak, M.P., Stolyar, S., Arkin, A.P., Leigh, J.A., Zhou, J., Keasling, J.D., Mukhopadhyay, A. and Stahl, D.A. (2012) Functional responses of methanogenic archaea to syntrophic growth. *ISME J*. **6**, 2045-2055.

Walsh, P.S., Erlich, H.A. and Higuchi, R. (1992) Preferential PCR amplification of alleles: mechanisms and solutions. *PCR Meth*. *Appl*. **1**, 241–250.

Werner, J.J., Zhou, D., Caporaso, G., Knight, R. and Angenent, L.T. (2011) Comparison of Illumina paired-end and single-direction sequencing for microbial 16S rRNA gene amplicon surveys. *ISME J*. **6**, 1273-1276.

Wilderer, P.A., Bungartz, H.J., Lemmer, H., Wagner, M., Keller, J. and Wuertz S. (2002) Modern scientific methods and their potential in wastewater science and technology. *Water Res*. **36**, 370-393.

Wilkie, A. and Colleran, E. (1989). The development of the anaerobic fixed-bed reactor and its application to the treatment of agricultural and industrial wastes. In: D.L. Wise (ed), International Biosystems, III. CRC Press, Inc., Boca Raton, FL, 183-226.

Wilkins, M.R., Sanchez, J.C., Gooley, A.A., Appel, R.D., Humphery-Smith, I., Hochstrasser, D.F. and Williams, K. L. (1995) Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. Biotech. *Gen*. *Eng*. *Rev*. **13**, 19-50.

Williams, T.J., Lauro, F.M., Ertan, H., Burg, D.W., Poljak, A., Raftery, M.J. and Cavicchioli, R. (2011) Defining the response of a microorganism to temperatures that span its complete growth temperature range (-2°C to 28°C) using multiplex quantitative proteomics. *Environm*. *Microbiol*. **13**(8), 2186 – 2203.

Williams, M.A., Taylor, E.B. and Mula, H.P. (2010) Metaproteomic characterization of a soil microbial community following carbon amendment. *Soil Biol*. *Biochem*. **42**, 1148–1156.

Wilmes, P. and Bond, P.L. (2004) The application of two- dimensional polyacrylamide gel electrophoresis and downstream analyses to a mixed community of prokaryotic microorganisms. *Environ*. *Microbiol*. **6**, 911–920.

Wilmes, P. and Bond, P.L. (2006) Metaproteomics: studying functional gene expression in microbial ecosystems. *Trends Microbiol*. **14**, 92–97.

Wilmes, P., Wexler, M. and Bond, P.L. (2008) Metaproteomics provides functional insight into activated sludge wastewater treatment. *PLoS ONE*. **3**, e1778.

von Wintzingerode, F., Göbel, U.B. and Stackebrandt, E. (1997) Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiol*. *Rev*. **21**, 213-229.

Wittwer, C.T., Herrmann, M.G., Moss, A.A. and Rasmussen, R.P. (1997) Continuous fluorescence monitoring of rapid cycle DNA amplification. *Biotechn*. **22**, 130–131.

Woese, C.R. (1987) Bacterial Evolution. *Microbiol*. *Rev*. **51**(2), 221-271.

Yakimov, M.M., Giuliano, L., Gentile, G., Crisafi, E., Chernikova, T.N., Abraham, W.R., Lünsdorf, H., Timmis, K.N. and Golyshin, P.N. (2003) *Oleispira antarctica* gen. nov., sp. nov., a novel hydrocarbonoclastic marine bacterium isolated from Antarctic coastal sea water. *Int*. *J*. *Syst*. *Evol*. *Microbiol*. **53**, 779–785.

Yu, Y., Lee, C., Kim, J. and Hwang, S. (2005) Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnol*. *Bioeng*. **89**, 670-679.

Zepp-Falz, K., Holliger, C., Grosskopf, R., Liesack, W., Nozhevnikova, A.N., Muller, B., Wehrli, B. and Hahn, D. (1999) Vertical distribution of methanogens in the anoxic sediment of Rotsee (Switzerland). *Appl*. *Environ*. *Microbiol*. **65**, 2402–2408.

Zhang, T. and Fang, H.H.P. (2006) Applications of real-time polymerase chain reaction for quantification of microorganisms in environmental samples. *Appl*. *Microbiol*. *Biotechnol*. **70**, 281–289.

Zhao, J.S., Manno, D., Leggiardo, C., O'Neil, D. and Hawari, J. (2006) *Shewanella halifaxensis* sp. nov., a novel obligately respiratory and denitrifying psychrophile. *Int*. *J*. *Syst*. *Evol*. *Microbiol*. **56**, 205–212.

Zhang, K., Martiny, A.C., Reppas, N.B., Barry, K.W., Malek, J., Chisholm, S.W., Church, G.M. (2006) Sequencing genomes from single cells by polymerase cloning. *Nat*. *Biotechnol*. **24**, 680–686.

Zhang, Y., Wen, Z., Washburn, M.P. and Florens, L. (2010) Refinements to Label Free Proteome Quantitation: How to Deal with Peptides Shared by Multiple Proteins. *Anal*. *Chem*. **82**(6), 2272-81.

Zhang, D., Zhu, W., Tang, C., Suo, Y., Gao, L., Yuan, X., Wang, X. and Cui, Z. (2012) Bioreactor performance and methanogenic population dynamics in a low-temperature (5–18 °C) anaerobic fixed-bed reactor. *Bioresour*. *Technol*. **104**, 136–143.

Zhu, J., Zheng, H., Guomin, A., Zhang, G., Liu, D., Xiaoli, L., Dong, X., 2012. The genome characteristics and predicted function of methyl-group oxidation pathway in the obligate acetoclastic methanogens, *Methanosaeta* spp. *PLoS ONE*. **7**(5), e367.

Chapter 2

A metaproteomic approach provides functional insights into microbial community dynamics during low-temperature anaerobic digestion

A condensed version of this chapter was submitted to Water Research:

Gunnigle, E., Siggins, A., Abram, F., Botting, C.H., O'Flaherty, V. (2012). A Metaproteomic approach provides functional insights into microbial community dynamics during low-temperature anaerobic digestion. Submitted Dec 2012.

Abstract

This study aimed to characterise the key microbial functional groups present in three laboratory-scale expanded granular sludge bed (EGSB) bioreactors; R1 (37°C), R2 (15°C) and R3 (7°C), which were employed to treat a volatile fatty acid-based wastewater (organic loading rates 3kg COD m^{-3} day^{-1} [R1, R2] and 3kg – 0.75kg COD m^{-3} day^{-1} [R3]). Metaproteomics, in conjunction with 16S rRNA gene phylogenetic approaches (clone libraries, qPCR), was applied to record microbial community composition and metaproteomic profiles as a function of bioreactor operating temperature. Clone libraries indicated a predominance of the *Chloroflexi* (21%) and *δ-Proteobacteria* (61%) bacterial groups in R1 sample, with *Firmicutes* (24%) and *Bacteroidetes (*46%) more prominent in R2 and R3 when compared with seed and R1 samples. The *Methanosaeta* genus was strongly represented in archaeal clone libraries, comprising 29% (R1), 76% (R2) and 91% (R3) of archaeal operational taxonomic units (OTUs) analysed. This was reflected in the metaproteomic results with 26 (65%) differentially expressed proteins assigned to this methanogen. Also evident from the metaproteomic data were proteins assigned to the bacterial phyla *Proteobacteria*, *Firmicutes* and *Actinobacteria*, while the archaeal orders *Methanobacteriales*, and *Methanomicrobiales* were also represented. The specific functions of the proteins identified related primarily to methanogenesis, with biosynthesis, cellular transport, energy, and cell wall maintenance proteins also evident. This study demonstrated that the metabolic diversity of *Methanosaeta* may allow it to adapt to lower temperatures, for example through expression of diverse acetyl CoA synthase genes. Furthermore, the possibility that a functionally significant group in low-temperature anaerobic digestion may be present in low numbers was suggested by the detection of a protein assigned to *Methanosarcina barkeri*, with this genus detected below the limit of quantification of the qPCR assay and no representative detected in clone library analysis. In overall terms, this study illustrated the value of employing a polyphasic approach to address microbial community dynamics in LTAD samples.

2.1 Introduction

Employing anaerobic digestion technology in a low-temperature context holds economic incentives over traditional mesophilic (>20°C) and thermophilic (>45°C) approaches through the reduced operation costs associated with treatment of dilute wastewaters (Enright *et al*., 2005; Connaughton *et al*., 2006; McKeown *et al*., 2009). Low-temperature anaerobic digestion (LTAD) has been successfully applied in laboratory-scale studies to treat a vast range of wastewater types, which include domestic (Cui *et al*., 2007), pharmaceutical (Enright *et al*., 2007), phenolic (Scully *et al*., 2006), chlorinated aliphatic (Siggins *et al*., 2011b), and brewery (Connaughton *et al*., 2006) -based wastewaters. Evidence of comparable treatment efficiencies to mesophilic counterparts has also been well documented (McHugh *et al*., 2004; Collins *et al*., 2005). The positive development towards the application of LTAD technology has been facilitated by efficient bioreactor designs such as the expanded granular sludge bed (EGSB) with the addition of positive modifications e.g. a fixed-film filter section (Collins *et al*., 2006), thus contributing to the successful adaptation of AD for the sub-mesophilic treatment of wastewater.

Despite successful applications, there is a lack of fundamental knowledge relating to the mechanisms underpinning AD. The design of bioreactors is generally based on rule of thumb, and bioreactor over-dimensioning; process instability and failures are still common. AD is operated based on relationships between bioreactor performance and empirical operating parameters, but the differences between successful and unsuccessful bioreactors are poorly understood. The future full-scale implementation of AD, and particularly the development of promising new applications, such as LTAD, is severely impaired by this knowledge gap. Methanogenic populations have been the focus of many LTAD studies due to their crucial role in biogas formation and biofilm integrity (Liu *et al*., 2002). Much of this work has focused on uncovering temporal methanogenic community dynamics under various operating temperatures primarily using nucleic acid based methods (O'Reilly *et al*., 2009; McKeown *et al*., 2009). In previous LTAD bioreactor trials, the prominence of the hydrogenotrophic *Methanomicrobiales* order has been recorded with McKeown *et al*. (2009), finding that a *Methanocorpusculum parvum*-like clone comprised 30% of

the archaeal clone library data, following a long-term LTAD trial. Also, the *Methanosaeataceae* family has been found to be a key methanogenic group underpinning LTAD bioreactor performance. This point was demonstrated in a study by O'Reilly *et al*. (2009), where high *Methanosaeataceae* detection (10^{11} copies 16S rRNA gene/$gVSS^{-1}$) correlated with a high chemical oxygen demand (COD) removal efficiency (>80%) at 15°C.

The above studies have provided the foundations from which targeted studies incorporating more sensitive approaches (e.g. RNA, protein and metabolite analysis) can be employed to elucidate functional information from complex systems such as LTAD bioreactors. Metaproteomics, which can be defined as the identification of all proteins expressed at a given time within an ecosystem (Wilmes & Bond, 2004), is an essential component of this function-based approach. Linking community structure (DNA based) to function (protein based) could provide a greater level of understanding of the LTAD process (Madden *et al*., 2010).

Abram *et al*. (2011) demonstrated the feasibility of applying a metaproteomic approach to LTAD, when examining the microbial community structure and protein profiles of granular sludge samples taken from a bioreactor treating a glucose based wastewater at 15°C. In order to investigate the physiology and adaptation mechanisms of methanogenic communities subjected to sub-mesophilic temperatures, the present study applied a polyphasic approach involving both metaproteomics and phylogenetic analyses to characterise the microbial communities from three independent anaerobic EGSB bioreactors operated at 37°C, 15°C and 7°C.

2.2 Materials and Methods

2.2.1 Source of biomass

Anaerobic granular sludge samples were obtained from three laboratory-scale EGSB bioreactors originally used as controls in LTAD studies (R2 [37°C], R4 [15°C]

Siggins *et al*., [2011a, b] and R2 [7°C] Siggins *et al*., [2011c]), which will be referred to as R1 (37°C), R2 (15°C) and R3 (7°C) for the remainder of the study. Each bioreactor had a working volume of 3.5 L. R1 and R2 were operated for 631 days (343 days [Siggins *et al*., 2011a], 288 days [Siggins *et al*., 2011b]), while R3 was operated for 609 days, 438 of these at 7°C after acclimation from 15°C (Siggins *et al*., 2011c). The bioreactors treated a synthetic volatile fatty acid (VFA) wastewater with the influent consisting of acetic acid, propionic acid, butyric acid and ethanol with a chemical oxygen demand (COD) ratio of 1:1:1:1, to a total of 3g COD L^{-1} (R1 & R2) and 3g - .75g COD L^{-1} (R3) buffered with $NaHCO_3$. The bioreactors had been originally seeded with anaerobic granular sludge from a full-scale (1500m^3) internal circulation (IC) anaerobic digester, operated at 37°C at the Carbery Milk Products plant (Ballineen, Co. Cork, Ireland). A concise summary of operation parameters as well as COD removal efficiencies and biogas production levels achieved during these bioreactor trials can be found in Table 2.1.

Table 2.1. Summary of R1 (37°C), R2 (15°C) and R3 (7°C) bioreactor operational performance adapted from Siggins *et al*, (2011 a, b, c). Values are means of phases, standard deviations are given in parenthesis.

Phases	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**	**P11**	**P12**	**P13**
Days (R1, R2)	0-149	150-172	173-191	192-226	227-243	244-342	343-390	391-464	465-484	485-518	519-539	540-574	575-631
Days (R3)	0-73	74-171	172-230	231-417	418-499	500-521	533-609						
Influent COD R1, R2 (Mgl^{-1})	3000	3000	3000	3000	3000	3000	3000	3000	3000	3000	3000	3000	3000
Influent COD R3 (Mgl^{-1})	3000	3000	1500	750	750	750	70						
COD Removal (%)[a]													
R1 (37°C)	78(13)	87(10)	89(7)	83(10)	76(9)	84(10)	85(6)	86(9)	87(9)	90(5)	87(7)	83(10)	88(8)
R2 (15°C)	61(12)	75(10)	74(11)	78(9)	79(10)	83(9)	87(7)	87(7)	90(4)	89(7)	87(6)	85(13)	89(6)
R3 (7°C)	54(17)	62(16)	72(15)	75(14)	83(10)	83(11)	86(11)						
CH_4 (%) in biogas[b]													
R1 (37°C)	50(7)	61(4)	71(3)	69(2)	70(3)	70(7)	71(6)	71(7)	79(3)	77(4)	76(4)	72(3)	73(3)
R2 (15°C)	56(15)	70(3)	74(2)	73(3)	76(3)	74(7)	77(4)	74(7)	74(5)	70(6)	71(3)	61(19)	76(3)
R3 (7°C)	68(6)	73(4)	69(12)	75(6)	70(11)	75(2)	73(7)						

[a] Total Chemical Oxygen Demand (COD) Removal Efficiency
[b] Specific Methanogenic Activity (ml CH_4 g[VSS]$^{-1}d^{-1}$)

2.2.2 Extraction of genomic DNA

Total genomic DNA was extracted from granular sludge biomass sampled at the conclusion of the trial using an automated nucleic acid extractor (Magtration 12GC, PSS Co., Chiba, Japan). Prior to extraction, granular biomass was finely crushed

using a mortar and pestle under liquid nitrogen, and resuspended in sterile double distilled water to a ratio of 1:4 (vol/vol). A 100 µl aliquot of the biomass suspension was loaded per extraction. Each extraction was performed in duplicate and the extracted DNA was eluted in Tris-HCl buffer (pH 8.0) and stored at -20 °C until use.

2.2.3 Clone library analysis of bacterial and archaeal 16S rRNA gene

Partial bacterial 16S rRNA gene sequences were amplified with forward primer 27F (5' – AGA GTT TGA TCC TGG CTC AG – 3'; DeLong, 1992) and reverse primer 1392R (5' ACG GGC GGT GTG TRC – 3'; Lane *et al*., 1985), while partial archaeal 16S rRNA gene sequences were amplified with forward primer 21F (5' – TTC CGG TTG ATC CYG CCG GA – 3'; Stackebrandt & Goodfellow, 1991) and reverse primer 958R (5' –YCC GGC GTT GAM TCC AAT T – 3'; DeLong, 1992). Primer details are in Appendix 1. Reaction mixtures (50 µl) contained 1.5 mM $MgCl_2$, 5 µl 10X NH_4 buffer (16 mM $(NH_4)2SO_4$, 67 mM Tris-HCl (pH 8.8 at 25°C), 0.01% Tween-20), 50 pmol dNTP (dATP, dCTP, dGTP, dTTP), 12.5 pmol of each primer, 200 ng template DNA and 0.2U Taq DNA polymerase. The PCR reactions were carried out using a touchdown PCR under the following conditions: denaturation at 95°C for 1 minute, annealing of primers (63°C – 54°C; 1 cycle at 1°C increments; 20 cycles at 52°C) for 1 minute and extension at 72°C for 2 minutes, followed by a final 10 minute extension at 72°C. Controls containing no DNA were also employed to identify amplification of contaminants. 5 µl of each PCR product was visualised after electrophoresis on 1% agarose TAE (0.5% w/v Tris, 0.11 % w/v acetic acid, 0.04% w/v EDTA) gel containing 1 µg ml-1 Sybr® Safe (Invitrogen), with Hyperlader IV (Bioline) as a molecular weight marker. Subsequent construction of clone libraries (TOPO®XL), amplified ribosomal DNA restriction analysis (ARDRA) and plasmid sequencing was performed as described by Collins *et al*. (2003). Any vector contamination was removed by screening sequence data using National Center for Biotechnology Information (NCBI) Vecscreen software. The resulting sequence data was compared to nucleotide databases using basic local alignment NCBI search tool (BLASTn) as described by Altschul *et al*. *(*1997). Sequences were aligned using MacClade 4.0 software (Sinauer Assoc) with nearest relatives from the BLASTn database and selected sequences downloaded from the

Ribosomal Database Project (RDP). Phylogenetic trees were constructed using the GTR + gamma model DNA substitution implemented RAxML 7.0.3 (Stamatakis, 2006) with all parameters optimized by RAxML. Confidence levels in the groupings of the phylogeny were assessed using 1000 bootstrap replicates as part of the RaXML phylogeny reconstruction. The resulting partial 16S rRNA gene sequences generated by this study were deposited in the Genbank database under the accession numbers (R1) HQ655412-HQ655420; (R2) HQ655421-HQ655434; (R3) HQ655435-HQ655457. In addition to gene sequences generated in this study, bacterial 16S rRNA gene sequences generated from seed biomass (Siggins, PhD thesis, NUI Galway, Ireland 2010), as well as archaeal sequences from seed (Siggins *et al*., 2011b) and end-point samples (Siggins *et al*., 2011c) were also used in order to give a concise overview of the community structure underpinning these bioreactors.

2.2.4 qPCR

Quantitative real-time PCR was performed using a LightCycler 480 (Roche, Mannheim, Germany). Four-methanogenic 16S rRNA gene primer and probe sets were used, specific for two hydrogenotrophic orders (*Methanomicrobiales* and *Methanobacteriales)* and two acetoclastic families (*Methanosaetaceae* and *Methanosarcinaceae*), accounting for most methanogens present in anaerobic digesters (Lee *et al*., 2009; Yu *et al*., 2005; Appendix 1). All DNA samples were analysed with each primer and probe set in duplicate. Each reaction mixture was prepared using the LightCycler TaqMan Master Kit (Roche): 2µl PCR-grad water, 1µl of probe (final concentration 200nM), 1µl each primer (final concentration 500nM), 10µl of 2X reaction solution and 5µl of DNA template. Amplification was carried out using a two-step thermal cycling protocol consisting of predenaturation for 10 minutes at 94°C, followed by 50 cycles of 10 seconds at 94°C and 30 seconds at 60°C.

Quantitative standard curves were constructed using the standard plasmids containing the full-length 16S rRNA gene sequences from the representative strains

of the target methanogenic groups as previously described (Lee *et al*., 2009; Yu *et al*., 2005). For each primer and probe set, an equimolar mixture of its corresponding standard plasmids was used as the template solution for constructing the standard curve. The mass concentration of each plasmid was measured in duplicate using a Qubit system (Invitrogen) and converted into its copy concentration as previously described (Lee *et al*., 2009). A 10-fold serial dilution series (10^1 -10^9 copies μl^{-1}) was generated for each standard solution and analyzed by real-time PCR in triplicate with its corresponding primer and probe set. The threshold cycle (CT) values determined were plotted against the logarithm of their input copy concentrations. The 16S rRNA gene copy concentrations of target groups were then estimated against the corresponding standard curves within the linear range (R2 > 0.995). The volume-based concentration (copies l^{-1}) were converted into the biomass-based concentration (copies g $[VSS]^{-1}$) using the VSS concentration of each sludge sample.

2.2.5 Two-dimensional gel electrophoresis

Proteins were extracted from 50 ml granular sludge samples from R1, R2 and R3 at the conclusion of the trial and subsequently separated by 2-DGE using a protocol as described by Abram *et al*. (2009). Briefly, the first dimension consisted of isoelectric focusing (IEF) using 7 cm IPG strips with linear pH gradients (pH 4 to 7; Amersham). The second dimension polyacrylamide (12% [w/v]) gels were run in pairs along with molecular weight markers with a range of 10-225 kDa (Broad Range Protein Molecular Markers, Promega). Gels were stained overnight in GelCode® 135 Blue staining reagent (Pierce) and then destained in deionised, distilled water for several hours. Gel images were captured by scanning with an Epson Perfection V350 photo scanner at a resolution of 800 dpi. Twenty-four gels were run corresponding to two duplicate independent extractions and four technical replicates of the three samples. Gel images were processed and analysed with PDQuest-Advanced software, version 8.0.1 (BioRad). Spot counts were obtained using the spot detection wizard enabling the Gaussian model option as recommended by the manufacturer. Ratios of spot intensities were determined for each bioreactor sample. Protein expression ratios greater than two-fold were considered significant.

Proteins of interest were excised from the gels and subjected to in-gel digestion prior to LC-MS/MS analysis.

2.2.6 Liquid chromatography and protein identification

Nanoflow liquid chromatography-electrospray ionization tandem mass spectrometry (nLC-ESI-MS/MS) and protein identification was carried out on proteins excised from the gels and subjected to in-gel digestion, using a ProGest Investigator in gel-digestion robot (Genomic Solutions) following standard protocols (Shevchenko *et al*., 1996). Briefly the gel protein samples were destained by washing with acetonitrile and subjected to reduction and alkylation before digestion with trypsin at 37°C. The peptides were extracted with 10% formic acid and concentrated down to 20 mL using a SpeedVac (ThermoSavant). They were then separated using an UltiMate nanoLC (LC Packings, Amsterdam) equipped with a PepMap C18 trap and column, using a 60 or 90 min elution profile, with a gradient of increasing acetonitrile, containing 0.1 % formic acid, to elute the peptides (5-35% acetonitrile in 18 (or 40 min for the 90 min elution) min, 35-50% in a further 7 (20 min) min, followed by 95% acetonitrile to clean the column, before requilibration to 5% acetonitrile). The eluent was sprayed into a Q-Star XL tandem mass spectrometer (ABSciex, Foster City, CA) and analysed in Information Dependent Acquisition (IDA) mode, performing 1 sec of MS followed by 3 sec MSMS analyses of the 2 most intense peaks seen by MS. These masses are then excluded from analysis for the next 60 sec. MS/MS data for +1 to +5 charged precursor ions which exceeded 150 cps was processed using the Paragon™ search algorithm (Shilova *et al*, 2007) within ProteinPilot 4.0 software (ABSciex, Foster City, CA) against NCBInr database July 2012 (38496380 sequences) with no species restriction.

2.3 Results

2.3.1 Clone library analysis

A total of 57 operational taxonomic units (OTUs) identified by screening 513 bacterial clones were analysed from seed and bioreactor biomass samples (Table 2.2). Bacterial OTUs relating to *δ-proteobacteria* (51%) were found to be predominant in the seed sludge (Fig. 2.1). Examples of notable shifts in community structure, which occurred during the trials, include a recorded increase in OTUs relating to the *Chloroflexi* (10% - 21%) group in R1 (operated at 37°C) by comparison to seed biomass sample. There was an increase in OTUs relating to *Firmicutes* (0% - 24%) recorded in R2 (operated at 15°C) and *Bacteroidetes* (15% - 46%) in R3 (operated at 7°C) (Fig. 2.1; Appendix 2).

There were 26 different OTUs identified in archaeal clone libraries (Table 2.3). Specifically, the seed biomass contained eight OTUs, R1 biomass also had eight, R2 had six and R3 contained four OTUs. Archaeal clones related to *Methanomicrobiales* were found to be dominant in the seed biomass (49%) with two OTUs closely aligned to *Methanolinea* species (Fig. 2.2). In this analysis, based on end-point PCR, *Methanomicrobiales*-like clones only represented 2% of R1 biomass, and were not detected in R3 biomass. However, a single OTU closely aligned to a *Methanocorpusculum* species comprised 19% of the R2 biomass sample (Fig. 2.2). *Methanobacteriales*-like clones were present in all biomass samples with a single OTU accounting for 30% of archaeal community in R1 biomass clone libraries. Although not detected in the seed biomass, *Methanosaeta*-like clones were the major constituent of the archaeal community in R2 (76%) and R3 (91%) biomass samples (Fig. 2.2). *Crenarchaeota*-like clones accounted for 24% of the archaeal community structure in the seed sample. However, their relative abundance was reduced in bioreactor samples with R1 having a single *Crenarchaeaota*-like OTU comprising 14% of the archaeal community, with no representatives detected in R2 or R3 biomass samples (Fig. 2.2).

Table 2.2. Phylogenetic affiliation of bacterial 16S rRNA gene sequences retrieved from biomass samples - seed sludge (day 0), R1 (37°C) and R2 (15°C) from day 631 and R3 (7°C) from day 609. The number of clones per OTU is given. Also, % similarity between cloned 16S rRNA gene and closest relatives in the NCBI database is indicated.

Library	Accession no.	No. of unique OUTs	Length (bp)	Closest relative (Blastn, Sequence Similarity %)	Phylogenic affiliation
Seed	HM749827	5	910	Uncultured *Bacteroidales* GU472718 (91%)	*Bacteroidetes*
Seed	HM749829	7	933	Uncultured bacterium GQ423770 (99%)	Unclassified bacteria
Seed	HM749830	5	817	Uncultured *Bacteroidetes* CU927222 (89%)	*Bacteroidetes*
Seed	HM749831	3	868	*Propionicimonas paludicola* FR733712 (97%)	*Actinobacterium*
Seed	HM749832	4	840	Uncultured *Aminanaerobia* CU918588 (99%)	*Synergistetes*
Seed	HM749833	47	695	Uncultured *Syntrophobacter* GU202953 (99%)	*δ- proteobacteria*
Seed	HM749834	3	865	Uncultured *Bacteroidetes* CU926283 (91%)	*Bacteroidetes*
Seed	HM749835	4	860	Uncultured *Chloroflexi* CU922816 (99%)	*Chloroflexi*
Seed	HM749836	2	978	Uncultured *Bacteroidetes* HQ183940 (94%)	*Bacteroidetes*
Seed	HM749837	4	907	Uncultured *Spirochaetes* CU925939 (98%)	*Spirochaete*
Seed	HM749838	3	930	Uncultured *Candidate division AC1* AY193177 (90%)	*δ- proteobacteria*
Seed	HM749839	1	845	Uncultured *Chloroflexi* HQ183907 (98%)	*Chloroflexi*
Seed	HM749840	3	926	Uncultured *Aminanaerobia* CU918588 (99%)	*Synergistetes*
Seed	HM749841	4	930	Uncultured *Chloroflexi* CU926360 (92%)	*Chloroflexi*
Seed	HM749842	2	855	Uncultured *Actinobacterium* HQ183929 (99%)	*Actinobacteria*
Seed	HM749843	1	843	Uncultured *Chloroflexi* JQ996665 (99%)	*Chloroflexi*
R1	HQ655412	39	865	Uncultured bacterium GQ181941 (99%)	*δ- proteobacteria*
R1	HQ655415	19	648	Uncultured *Chloroflexi* HQ132968 (99%)	*Chloroflexi*
R1	HQ655416	12	800	*Uncultured Sytrophomonas* EU887790 (97%)	*Firmicutes*
R1	HQ655417	6	786	Uncultured *Chloroflexi* CU922521 (99%)	*Chloroflexi*

Table 2.2. Continued

Library	Accession no.	No. of unique OUTs	Length (bp)	Closest relative (Blastn, Sequence Similarity %)	Phylogenic affiliation
R1	HQ655418	56	852	Uncultured *Syntrophaceae* GU202942 (96%)	*δ- proteobacteria*
R1	HQ655419	8	769	*Levilinea saccharolytica* NR040972 (100%)	*Chloroflexi*
R1	HQ655420	15	854	Uncultured *Cytophagales* FJ516908 (99%)	*Bacteroidetes*
R2	HQ655421	5	865	Uncultured *Dehalobacter* HM748813 (100%)	*Firmicutes*
R2	HQ655422	13	648	Uncultured *Prolixibacter* JQ012279 (99%)	*Bacteroidetes*
R2	HQ655423	5	813	*Azospirillum brasilense* FR667907 (90%)	*α proteobacteria*
R2	HQ655424	18	825	Uncultured *Syntrophomonas* JQ599687 (98%)	*Firmicutes*
R2	HQ655425	13	533	Uncultured *Chloroflexi* JN038747 (91%)	*Chloroflexi*
R2	HQ655426	30	697	Uncultured *Candidatus cloacamonas* EU887773 (97%)	Unclassified bacteria
R2	HQ655427	4	880	Uncultured *Bacteroidetes* JQ724326 (99%)	*Bacteroidetes*
R2	HQ655429	6	813	Uncultured *Smithella* EU888819 (99%)	*δ- proteobacteria*
R2	HQ655430	9	586	Uncultured bacterium GQ181941 (99%)	*δ- proteobacteria*
R2	HQ655431	4	798	Uncultured *Pelotomaculum* AY607115 (96%)	*Firmicutes*
R2	HQ655432	1	898	Uncultured *Bacteroidetes* JF305757 (95%)	*Bacteroidetes*
R2	HQ655433	4	875	Uncultured *Syntrophorhabdaceae* HQ003592 (99%)	*δ- proteobacteria*
R2	HQ655434	2	768	Uncultured *Desulfobacteraceae* HQ133033 (99%)	*δ- proteobacteria*
R3	HQ655435	21	705	Uncultured *Bacteroidetes* CU926896 (99%)	*Bacteroidetes*
R3	HQ655436	18	905	Uncultured *Chloroflexi* CU922816 (99%)	*Chloroflexi*
R3	HQ655437	11	904	Uncultured *Syntrophus* GU112190 (99%)	*δ- proteobacteria*
R3	HQ655438	15	791	Uncultured *Bacteroidetes* JF305757 (99%)	*Bacteroidetes*
R3	HQ655439	2	909	Uncultured *Bacteriodetes* GQ406172 (97%)	*Bacteroidetes*

Table 2.2. Continued

Library	Accession no.	No. of unique OUTs	Length (bp)	Closest relative (Blastn, Sequence Similarity %)	Phylogenic affiliation
R3	HQ655440	4	851	Uncultured *Syntrophorhabdaceae* JF946911 (99%)	*δ- proteobacteria*
R3	HQ655441	1	919	Uncultured *Planctomycete* GQ356196 (95%)	*Planctomycetes*
R3	HQ655442	2	896	Uncultured *Bacteroidetes* FJ484674 (99%)	*Bacteroidetes*
R3	HQ655443	16	906	Uncultured *Firmicutes* CU919858 (99%)	*Firmicutes*
R3	HQ655444	10	930	Uncultured *Bacteroidetes* EU721793 (99%)	*Bacteroidetes*
R3	HQ655445	1	741	Uncultured *Aminanaerobia* CU926332 (100%)	*Synergistetes*
R3	HQ655446	1	891	Uncultured *Geobacter* AY607119 (96%)	*δ- proteobacteria*
R3	HQ655447	11	874	Uncultured *Bacteroidetes* HQ003601 (97%)	*Bacteroidetes*
R3	HQ655448	8	900	Uncultured *Planctomycete* (89%)	*Planctomycetes*
R3	HQ655449	1	850	*Geobacter hephaestius* AY737507 (97%)	*δ- proteobacteria*
R3	HQ655450	5	958	Uncultured *Bacteroidetes* JQ599633 (99%)	*Bacteroidetes*
R3	HQ655451	9	911	Uncultured *Pelotomaculum* AY607115 (97%)	*Firmicutes*
R3	HQ655453	1	917	Uncultured *Bacteroidetes* JQ012263 (98%)	*Bacteroidetes*
R3	HQ655455	1	891	Uncultured *Smithella* EU888819 (99%)	*δ- proteobacteria*
R3	HQ655456	2	919	Uncultured *Bacteroidetes* HQ183938 (99%)	*Bacteroidetes*
R3	HQ655457	6	891	Uncultured *candidate division TM7* GU180006 (93%)	*TM7*

Table 2.3. Phylogenetic affiliation of archaeal 16S rRNA gene sequences retrieved from biomass samples - seed sludge (day 0), R1 (37°C) and R2 (15°C) from day 631 and R3 (7°C) from day 609. The number of clones per OTU is given. Also, % similarity between cloned 16S rRNA gene and closest relatives in the NCBI database is indicated

Library	Accession no.	No. of unique OUT's	Length (bp)	Closest relative (Blastn, Sequence Similarity %)	Phylogenic affiliation
Seed	HM749805	24	745	Uncultured *Methanolinea sp.* AB479406 (97%)	*Methanomicrobiales*
Seed	HM749806	7	800	Uncultured *Crenarchaeota* CU916928 (99%)	*Crenarchaeotes*
Seed	HM749807	2	793	Uncultured *Crenarchaeota* U59986 (99%)	*Crenarchaeotes*
Seed	HM749808	11	792	Uncultured *Thermoplasmatales* EU731598 (95%)	*Thermoplasmatales*
Seed	HM749809	1	789	*Uncultured Methanobacterium* AB598270 (99%)	*Methanobacteriales*
Seed	HM749810	1	790	Uncultured *Methanomethylovorans* DQ631887 (99%)	*Methanosarcinales*
Seed	HM749811	5	792	Uncultured *Crenarchaeota* CU916928 (99%)	*Crenarchaeota*
Seed	HM749813	2	779	Uncultured *Methanolinea* AB479407 (99%)	*Methanomicrobiales*
R1	HQ434532	14	883	*Methanobacterium beijengense* AY552778 (99%)	*Methanobacteriales*
R1	HQ434533	9	880	Uncultured *Methanosarcinales* CU916680 (99%)	*Methanosarcinales*
R1	HQ434534	9	901	Uncultured archaeon GU196162 (99%)	*Uncultured Euryarchaeota*
R1	HQ434535	2	898	*Uncultured Methanosaeta* JQ766949 (99%)	*Methanosaeta*
R1	HQ434536	3	900	Uncultured archaeon GU196187 (99%)	*Uncultured Euryarchaeota*
R1	HQ434537	3	905	Uncultured *Methanosaeta* JN052765 (100%)	*Methanosaeta*
R1	HQ434538	6	902	Uncultured *Crenarchaeota* U59986 (95%)	*Crenarchaeota*
R1	HQ434539	1	909	Uncultured *Methanolinea* JQ668634 (99%)	*Methanomicrobiales*
R2	HQ434544	20	892	Uncultured *Methanosaeta* JN052768 (99%)	*Methanosaeta*
R2	HQ434545	12	890	*Methanocorpusculum sinense* FR749947 (99%)	*Methanomicrobiales*
R2	HQ434546	28	900	Uncultured *Methanosaeta* JN052771 (99%)	*Methanosaeta*
R2	HQ434547	1	901	*Methanobacterium beijengense* AY552778 (99%)	*Methanobacteriales*

Table 2.3. Continued

Library	Accession no.	No. of unique OUT's	Length (bp)	Closest relative (Blastn, Sequence Similarity %)	Phylogenic affiliation
R2	HQ434548	1	887	*Methanobacterium beijengense* AY552778 (99%)	*Methanobacteriales*
R2	HQ434549	1	899	Uncultured archaeon JQ668625 (99%)	*Uncultured Euryarchaeota*
R3	HQ315898	41	898	Uncultured *Methanosaeta* EU888812 (99%)	*Methanosaeta*
R3	HQ315899	3	887	*Methanobacterium beijengense* AY552778 (99%)	*Methanobacteriales*
R3	HQ315900	1	899	Uncultured *Methanosaeta* GU475190 (100%)	*Methanosaeta*
R3	HQ315901	1	879	Uncultured *Methanosaeta* JN052768 (99%)	*Methanosaeta*

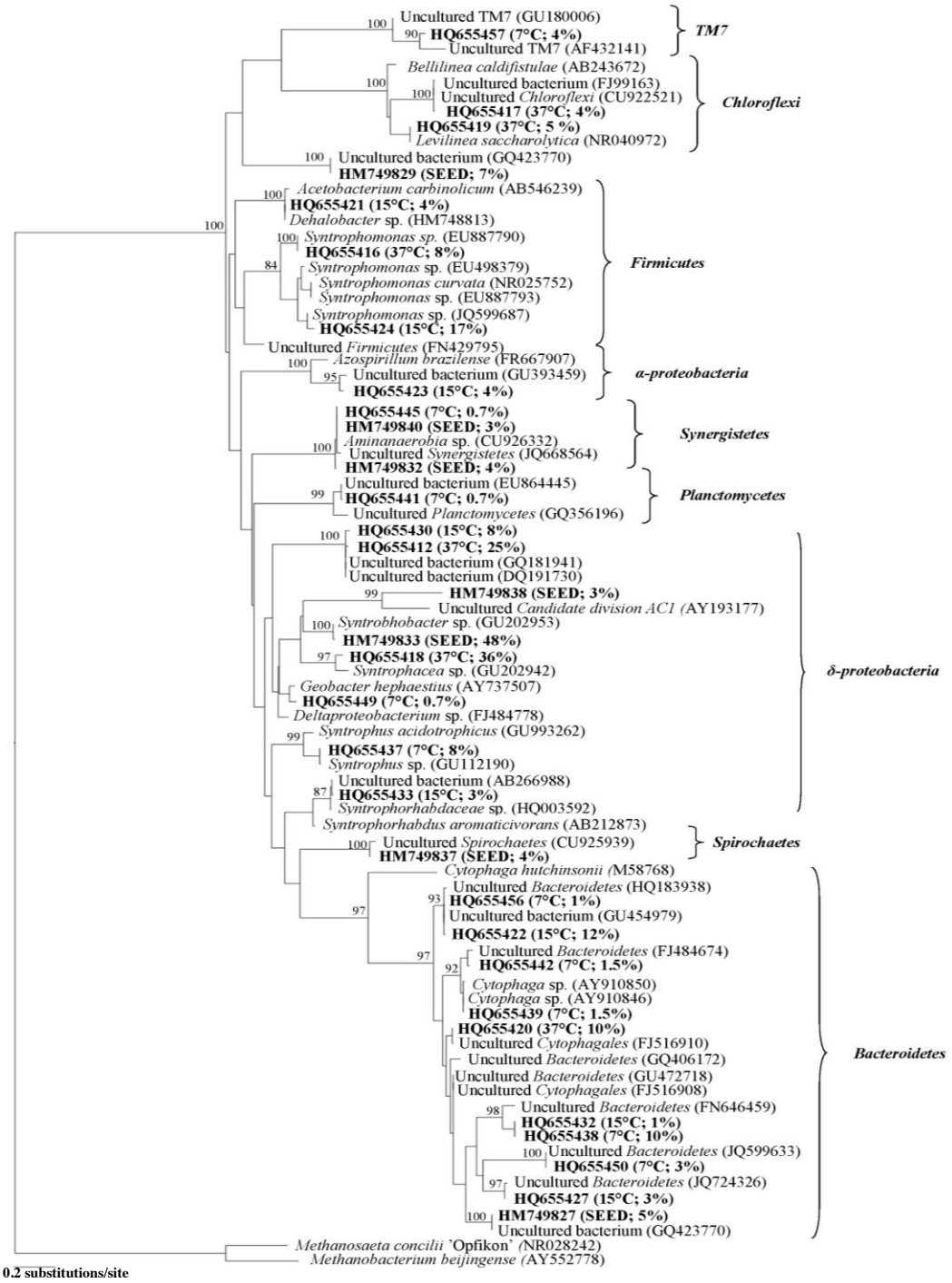
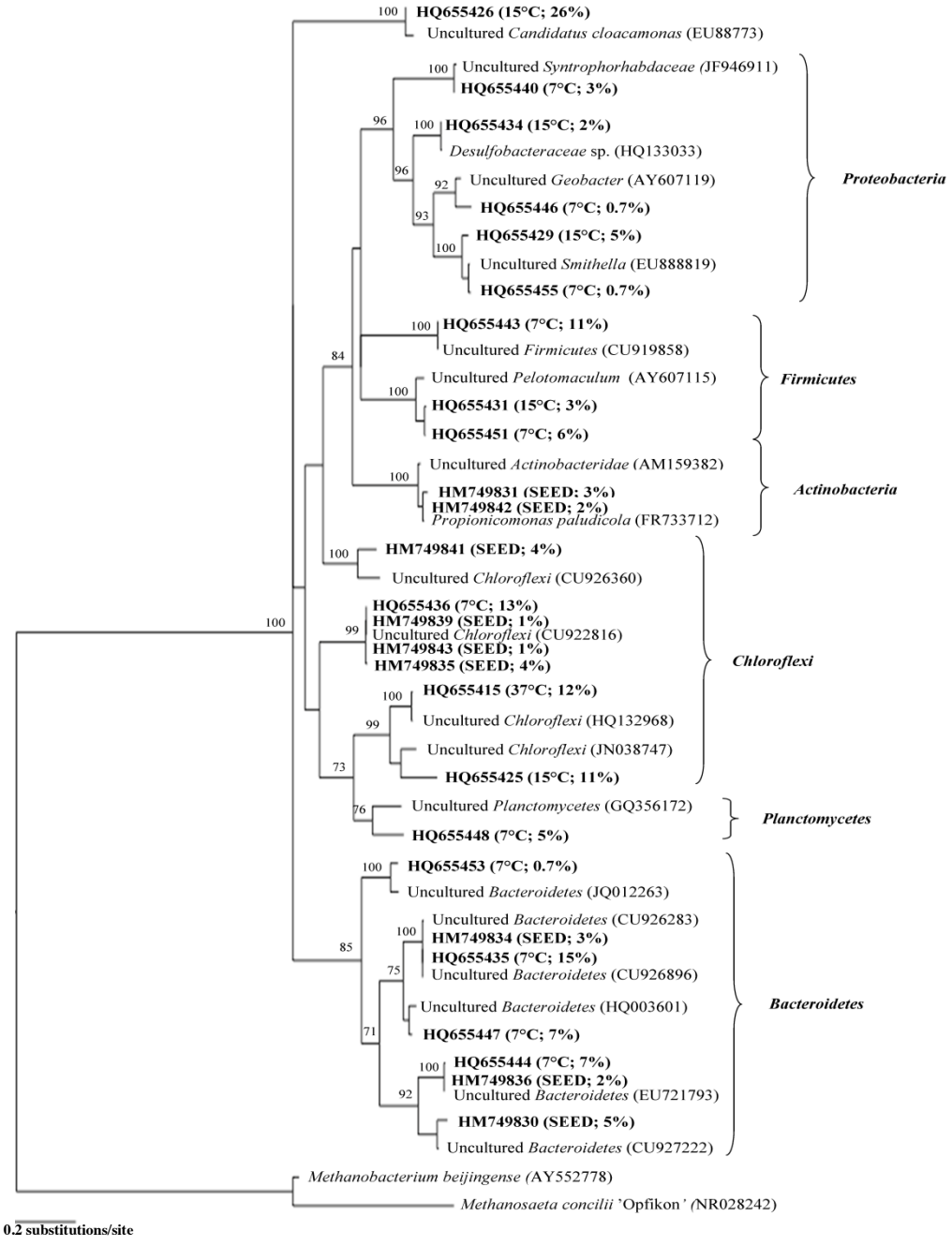


Figure 2.1 Phylogenetic affiliation of the bacterial 16S rRNA gene sequences identified from day 0 (seed biomass), day 631 (R1-37°C, R2-15°C) and day 609 (R3-7°C) bioreactor biomass, calculated using the GTR + gamma model DNA substitution implemented RAxML 7.0.3 (Stamatakis, 2006). Bootstrap replicates (total 1000 replicate samples) supporting the branching order are shown at relevant nodes.


HQ655426 (15°C; 26%)
Uncultured Candidatus cloacamonas (EU88773)
Uncultured Syntrophorhabdaceae (JF946911)
HQ655440 (7°C; 3%)
HQ655434 (15°C; 2%)
Desulfobacteraceae sp. (HQ133033)
Uncultured Geobacter (AY607119)
HQ655446 (7°C; 0.7%)
HQ655429 (15°C; 5%)
Uncultured Smithella (EU888819)
HQ655455 (7°C; 0.7%)
Proteobacteria
HQ655443 (7°C; 11%)
Uncultured Firmicutes (CU919858)
Uncultured Pelotomaculum (AY607115)
HQ655431 (15°C; 3%)
HQ655451 (7°C; 6%)
Firmicutes
Uncultured Actinobacteridae (AM159382)
HM749831 (SEED; 3%)
HM749842 (SEED; 2%)
Propionicomonas paludicola (FR733712)
Actinobacteria
HM749841 (SEED; 4%)
Uncultured Chloroflexi (CU926360)
HQ655436 (7°C; 13%)
HM749839 (SEED; 1%)
Uncultured Chloroflexi (CU922816)
HM749843 (SEED; 1%)
HM749835 (SEED; 4%)
HQ655415 (37°C; 12%)
Uncultured Chloroflexi (HQ132968)
Uncultured Chloroflexi (JN038747)
HQ655425 (15°C; 11%)
Chloroflexi
Uncultured Planctomycetes (GQ356172)
HQ655448 (7°C; 5%)
Planctomycetes
HQ655453 (7°C; 0.7%)
Uncultured Bacteroidetes (JQ012263)
Uncultured Bacteroidetes (CU926283)
HM749834 (SEED; 3%)
HQ655435 (7°C; 15%)
Uncultured Bacteroidetes (CU926896)
Uncultured Bacteroidetes (HQ003601)
HQ655447 (7°C; 7%)
HQ655444 (7°C; 7%)
HM749836 (SEED; 2%)
Uncultured Bacteroidetes (EU721793)
HM749830 (SEED; 5%)
Uncultured Bacteroidetes (CU927222)
Bacteroidetes
Methanobacterium beijingense (AY552778)
Methanosaeta concilii 'Opfikon' (NR028242)
0.2 substitutions/site

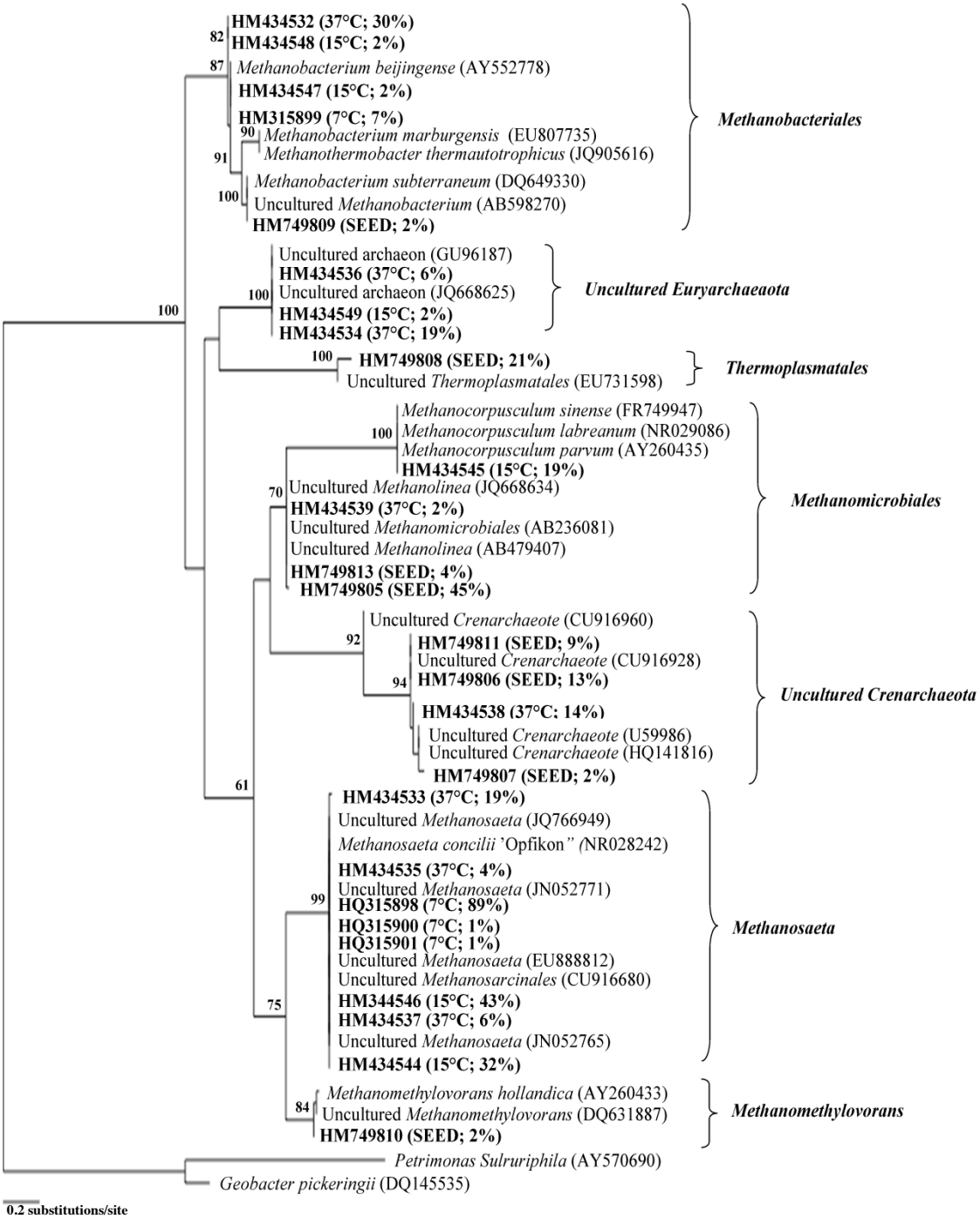


Figure 2.2 Phylogenetic affiliation of the archaeal 16S rRNA gene sequences identified from day 0 (seed biomass), day 631 (R1-37°C, R2-15°C) and day 609 (R3-7°C) bioreactor biomass, calculated using the GTR + gamma model DNA substitution implemented RAxML 7.0.3 (Stamatakis, 2006). Bootstrap replicates (total 1000 replicate samples) supporting the branching order are shown at relevant nodes.

2.3.2 qPCR

The qPCR results highlighted clear changes in the quantitative composition of the methanogenic community in each bioreactor. The hydrogenotrophic order *Methanobacteriales* was detected at consistently high levels, with each comprising $>10^9$ copies $gVSS^{-1}$ and was the highest recorded group in R1 biomass (Fig. 2.3), accounting for 70% of the total measured methanogenic population, i.e., the sum of all 16S rRNA gene concentrations quantified with the primer/probe sets used for this sample above the detection limit of 10^6 copies $gVSS^{-1}$.

The other hydrogenotophic order, *Methanomicrobiales*, displayed a greater level of fluctuation between bioreactor samples with a 2-log increase in R2 biomass when compared with R1 and R3 biomass samples (Fig. 2.3). However, the *Methanomicrobiales* (MMB) group had lower gene copy numbers than the *Methanobacteriales* (MBT) and *Methanosaeta* groups (MSt) with the highest MMB concentration of 1.23×10^9 copies $gVSS^{-1}$ recorded in R2 biomass (Fig. 2.3).

The acetoclastic *Methanosaetacaea* followed a similar trend to the hydrogenotrophic *Methanobacteriales* with all biomass samples containing $>10^9$ copies $gVSS^{-1}$ (Fig. 3). *Methanosaetacaea* was the dominant group in R2 biomass, representing 69% of the total measured methanogenic population in this sample, equating to 6.92×10^9 copies $gVSS^{-1}$ (Fig. 2.3).

The *Methanosarcinaceae* were only detected below the level of quantification in qPCR assay (10^6 copies $gVSS^{-1}$) and they were thus not included in the comparative analysis.

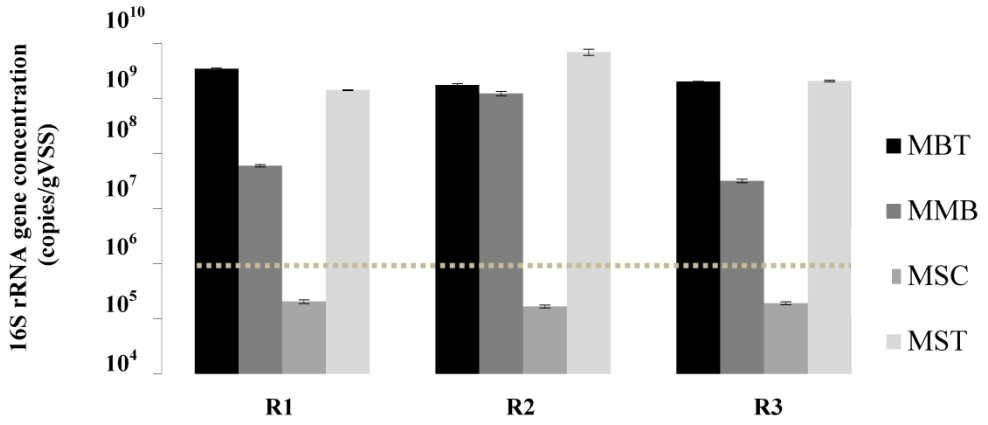


Figure 2.3. Absolute quantification of the 16S rRNA gene concentration of methanogens: **MBT** *Methanobacteriales*; **MMB** *Methanomicrobiales*; **MSC** *Methanosarcinicaeae*; **MST** *Methanosaetaceae* from biomass sampled on day 631 for R1 (37°C) and R2 (15°C) and day 609 for R3 (7°C). Error bars indicate the standard deviation and are the result of two replicates. Dashed line relates to detection limit of assay (10^6)

2.3.3 Metaproteomics

An average of 267 (SD 26.4; n=9) reproducible protein spots were detected on 2-D gels selected for PDQuest analysis. Sixty-five distinct spots were excised for protein identification resulting in 91% positive identifications using nLC-ESI-MS/MS. There were, however, a number of proteins that migrated as several distinct spots, which is well documented in relation to the 2-DGE technique (O'Farrell, 1975; Gygi *et al*., 2000). Thus, a final total of 42 proteins were identified (Table 2.4).

Overall, 37 proteins were assigned to the archaeal group, which included 9 distinct species belonging to the methanogenic orders *Methanobacteriales*, *Methanomicrobiales* and *Methanosarcinales*. Most of these proteins were found to be involved in methanogenesis from acetate (acetyl-CoA synthetase) and CO_2 (coenzyme F_{420} dependent N_5, N_{10} H_4MPT reductase; Table 2.4, Fig. 2.4). The acetyl-CoA synthetase (ACS) protein catalyses the activation of acetate to acetyl-CoA and was assigned to *Methanosaeta concilii*. Interestingly, this protein was present in the dataset relating to two separate gene loci (Table 2.4). The ACS protein relating to *mcon*_0558 gene was up regulated in low-temperature biomass with ACS protein relating to *mcon*_0559 gene was expressed at a higher level in mesophilic

biomass (Table 2.4). A bifunctional formaldehyde activating enzyme/3-hexulose-6-phosphate synthase (Fae/HPS) also assigned to *M. concilii* was detected at a higher level in both low-temperature biomass samples compared to mesophilic biomass (Table 2.4). The HPS protein was also identified separately from the Fae protein complex. This protein was once again assigned to *M. concilii* and was conserved for each biomass sample (Table 2.4). In addition, proteins involved in glycolysis, amino acid biosynthesis, solute transportation, protein repair and regulation were also identified and assigned to archaeal species (Table 2.4, Fig. 2.4).

A total of 5 proteins were assigned to bacterial species in this study. Three of these were assigned to the *Proteobacteria* phylum, with the *Firmicutes* and *Actinobacteria* phyla represented by a single protein each (Table 2.4). The functional categories of the bacterial proteins varied from cell wall maintenance to membrane transport and energy. Included was a GroEL chaperone assigned to a *Geobacter bemidjensis*, which was detected at a higher level in mesophilic biomass (R1) than at lower temperatures (R2 & R3; Table 2.4). Also identified was a LPXTG-motif cell wall anchor domain protein involved in cell wall maintenance and was significantly expressed at a higher level in R3 biomass sample compared to both R1 and R2 (Table 2.4). Finally, a host of hypothetical proteins of unknown functions were apparent in this dataset assigned to *M. concilii* and *Stigmatella aurantiaca*, a member of the *δ-proteobactyeria* phylum (Table 2.4).

Table 2.4 Identification and putative function assigned to proteins excised from 2D-gels derived from bioreactor biomass. Ratios relate to differential abundance of protein for the three biomass samples 37°C (R1), 15°C (R2) and 7°C (R3). Significantly expressed proteins are in bold. [a] corresponds with ACS protein related to *acs*2 gene locus. [b] corresponds with ACS protein related to *acs*3 gene locus

Protein name	Gene locus	Species assignment	Phylogenetic classification	Ratio R1/R2	Ratio R1/R3	Ratio R2/R3	Suggested function
5,10-methylenetetrahydro-methanopterin reductase	Mswan_1889	*Methanobacterium sp.*	Order *Methanobacteriales*	**2.9**	**2.3**	1.2	Methanogenesis from CO_2
3-hexulose-6-phosphate synthase	Mcon_0429	*Methanosaeta concilii*	Order *Methanosarcinales*	0.63	1	1	Biosynthesis
3-isopropylmalate dehydratase large subunit	Mcon_1320	*Methanosaeta concilii*	Order *Methanosarcinales*	**2**	0.61	0.69	Biosynthesis
Acetyl-coenzyme A synthetase[a]	Mcon_0558	*Methanosaeta concilii*	Order *Methanosarcinales*	**0.38**	**0.22**	1.7	Methanogenesis from acetate
Acetyl-coenzyme A synthetase[b]	Mcon_0559	*Methanosaeta concilii*	Order *Methanosarcinales*	**15.8**	**8.5**	1.2	Methanogenesis from acetate
Aspartate-semialdehyde dehydrogenase	Mcon_1211	*Methanosaeta concilii*	Order *Methanosarcinales*	**2.3**	1	1.6	Biosynthesis
Bifunctional enzyme Fae/Hps	Mcon_1383	*Methanosaeta concilii*	Order *Methanosarcinales*	**0.39**	**0.38**	**0.34**	Formaldehyde assimilation
Chaperonin GroEL	Mcon_0249	*Geobacter bemidjiensis*	Phylum *Proteobacteria*	**2.5**	**2.9**	1.6	Protein folding

Table 2.4 Continued

Protein name	Gene locus	Species assignment	Phylogenetic classification	Ratio R1/R2	Ratio R1/R3	Ratio R2/R3	Suggested function
Chaperone protein DnaK	Mcon_1037	*Methanosaeta concilii*	Order *Methanosarcinales*	**4.5**	**2.4**	**0.29**	Protein folding
CO dehydrogenase/acetyl-CoA synthase complex α-subunit	Mcon_1332	*Methanosaeta concilii*	Order *Methanosarcinales*	1.5	1.8	1.9	Methanogenesis from acetate
CO dehydrogenase/acetyl-CoA synthase β-subunit	Mcon_1330	*Methanosaeta concilii*	Order *Methanosarcinales*	1.2	1.3	0.64	Methanogenesis from acetate
CO dehydrogenase/acetyl-CoA synthase δ-subunit	Mcon_1326	*Methanosaeta concilii*	Order *Methanosarcinales*	0.81	0.73	0.64	Methanogenesis from acetate
Enolase	Mcon_0157	*Methanosaeta concilii*	Order *Methanosarcinales*	0.52	0.78	1.8	Glycolysis
Family 5 extracellular solute-binding protein	Mcon_2385	*Methanosaeta concilii*	Order *Methanosarcinales*	1.1	0.52	1.3	Transportation
Hypothetical protein MCON_2188	Mcon_2188	*Methanosaeta concilii*	Order *Methanosarcinales*	0.50	0.72	0.72	Unknown
Hypothetical protein STAUR_6660	Staur_6660	*Stigmatella aurantiaca*	Phylum *Proteobacteria*	0.50	1.2	1.6	Unknown
Hypothetical protein MCON_3417	Mcon_3417	*Methanosaeta concilii*	Order *Methanosarcinales*	**0.30**	**0.39**	1.8	Unknown

Table 2.4 Continued

Protein name	Gene locus	Species assignment	Phylogenetic classification	Ratio R1/R2	Ratio R1/R3	Ratio R2/R3	Suggested function
Hypothetical protein MCON_1190	Mcon_1190	*Methanosaeta concilii*	Order *Methanosarcinales*	**0.47**	**0.28**	1.4	Unknown
LPXTG-motif cell wall anchor domain protein	Spars_1364	*Streptococcus pneumoniae*	Phylum *Firmicutes*	1.9	**0.34**	**0.27**	Cell wall maintenance
Manganese-dependent inorganic pyrophosphatase	Mcon_1906	*Methanosaeta concilii*	Order *Methanosarcinales*	1	1.9	1.8	Cellular energy
Methyl-coenzyme M reductase I α-subunit	Mswan_2056	*Methanobacterium subterraneum*	Order *Methanobacteriales*	1.6	1	1.4	Methanogenesis
Methyl-coenzyme M reductase I α-subunit	Mcon_0759	*Methanosaeta concilii*	Order *Methanosarcinales*	**0.33**	1.3	**12.1**	Methanogenesis
Methyl-coenzyme M reductase I α-subunit	Mth_1164	*Methanothermobacter thermautotrophicus*	Order *Methanobacteriales*	**0.45**	1.8	**2.6**	Methanogenesis
Methyl-coenzyme M reductase I β-subunit	Memar_0617	*Methanoculleus marisnigri*	Order *Methanomicrobiales*	0.71	**18.6**	**11.7**	Methanogenesis
Methyl coenzyme M reductase I β-subunit	Mth_1132	*Methanothermobacter thermautotrophicus*	Order *Methanobacteriales*	1.9	0.80	0.65	Methanogenesis
Methyl coenzyme M reductase I β-subunit	Mhun_2144	*Methanospirillum hungatei*	Order *Methanomicrobiales*	0.72	**0.42**	**0.34**	Methanogenesis

Table 2.4 Continued

Protein name	Gene locus	Species assignment	Phylogenetic classification	Ratio R1/R2	Ratio R1/R3	Ratio R2/R3	Suggested function
Methyl coenzyme M reductase I β-subunit	Mcon_0762	*Methanosaeta concilii*	Order *Methanosarcinales*	**0.26**	1.1	**7.5**	Methanogenesis
Methyl-coenzyme M reductase I γ-subunit	Mcon_0760	*Methanosaeta concilii*	Order *Methanosarcinales*	**0.08**	0.82	**9.4**	Methanogenesis
Methyl-coenzyme M reductase I γ-subunit	Mtbma_c15490	*Methanothermobacter marburgensis*	Order *Methanobacteriales*	**43.9**	**6.5**	**0.07**	Methanogenesis
Periplasmic solute binding protein	Srot_0197	*Segniliparus rotundus*	Phylum *Actinobacteria*	**4.1**	**3.9**	0.80	Membrane transport
PIN domain-containing protein	Mcon_1569	*Methanosaeta concilii*	Order *Methanosarcinales*	**0.29**	1.3	**12.1**	Ribonuclease activity
Proteasome α-subunit	Mbar_a2503	*Methanosarcina barkeri*	Order *Methanosarcinales*	**0.40**	0.82	1.1	Cell regulation
Proteasome α-subunit	Mthe_1365	*Methanosaeta thermophila*	Order *Methanosarcinales*	1.8	1.6	0.83	Cell regulation
Putative methanogenesis marker protein 15	Mcon_0397	*Methanosaeta concilii*	Order *Methanosarcinales*	**0.44**	1.8	**2.1**	Methanogenesis
S-layer-related duplication domain-containing protein	Mcon_1153	*Methanosaeta concilii*	Order *Methanosarcinales*	1.2	0.81	1.3	Cell envelope maintenance

Table 2.4 Continued

Protein name	Gene locus	Species assignment	Phylogenetic classification	Ratio R1/R2	Ratio R1/R3	Ratio R2/R3	Suggested function
S-adenosylmethionine synthetase	Mcon_2824	*Methanosaeta concilii*	Order *Methanosarcinales*	**0.45**	1.8	**2.2**	Biosynthesis
Succinyl-CoA synthetase -subunit	Sfum_1703	*Syntrophobacter fumaroxidans*	Phylum *Proteobacteria*	0.78	0.83	1.6	Cellular energy
Tetrahydromethanopterin S-methyltransferase subunit H	Mcon_1068	*Methanosaeta concilii*	Order *Methanosarcinales*	**3.6**	1.3	**0.31**	Methanogenesis from CO_2
Thermosome α-subunit	Mcon_0154	*Methanosaeta concilii*	Order *Methanosarcinales*	**0.30**	**0.20**	0.83	Molecular chaperone
Thermosome δ-subunit	Mcon_1442	*Methanosaeta concilii*	Order *Methanosarcinales*	0.95	1.5	0.83	Molecular chaperone
V-type ATP synthase α-subunit	Mcon_2516	*Methanosaeta concilii*	Order *Methanosarcinales*	1.8	0.67	0.67	ATP synthesis
V-type ATP synthase β-subunit	Mcon_2515	*Methanosaeta concilii*	Order *Methanosarcinales*	1.7	1.7	1.2	ATP synthesis

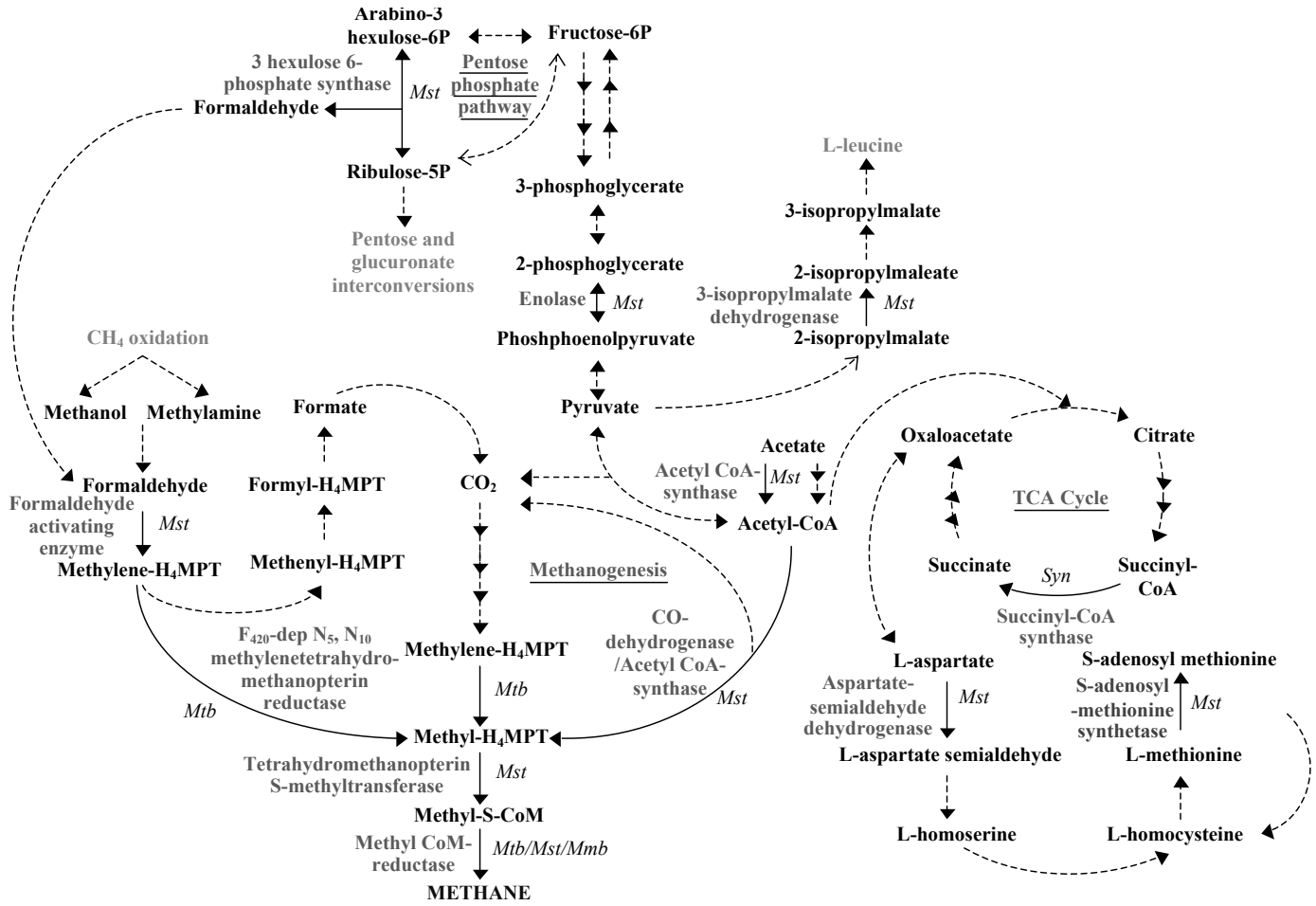


Figure 2.4 Proposed metabolic model for anaerobic digestion of synthetic volatile fatty acid wastewater inferred from metaproteomic data. Proteins identified in this study are located left of function undertaken. Dashed lines indicate steps of pathway not represented in these results. Not all intermediate metabolites are shown. *Mst*: *Methanosaeta*; *Mtb*: *Methanobacteriales*; *Mmb*: *Methanomicrobiales*; *Syn*: *Syntrophobacter*.

2.4 Discussion

The ability of microbial communities to adapt and achieve treatment efficiencies in low-temperature AD systems comparable to mesophilic setups has been demonstrated in a number of recent laboratory- and pilot-scale studies (O'Reilly *et al*., 2009; Siggins *et al*., 2011a, b; Martinez-Sosa *et al*., 2011; Zhang *et al*., 2012). The emergence of specific psychrophilic organisms may not be necessary to allow such successful LTAD treatment efficiencies and, indeed, mesophilic seed biomass appears to have pre-existing psychrotolerant populations (O'Reilly *et al*. 2009). It has also been reported, however, that truly psychrophilic organisms may emerge during long-term LTAD (McKeown *et al*., 2009). Further supporting evidence regarding the potential for successful LTAD was provided by the performance of two low-temperature bioreactors R2 (15°C) and R3 (7°C), compared to that of the mesophilic R1 (37°C) bioreactor, operated by Siggins *et al*. (2011a, b; Table 2.1) in trials of >600 days. In this study, clear differences in both the microbial community structure and protein expression levels were identified in these bioreactor biomass samples (Table 2.4, Figs. 2.1, 2.2).

In total, 42 proteins excised from 2D-gels were identified, 19 of which related to metabolism (Table 2.4). Methanogenesis was the primary component of this functional category with 17 proteins relating to this metabolic process. This result was comparable with a study by Abram *et al*. (2011), where one third of proteins identified related to methanogenesis. The relatively high proportion of proteins involved in methanogenesis suggests that the methanogenic populations were efficiently adapted to operation temperature, which was reflected by the high level of CH_4 production recorded in each bioreactor at end of trials ≥ 70% (Table 2.1). Moreover, due to the simple nature of the VFA wastewater used in this study, the finding of significant functional activity relating to methanogenic species was not surprising.

Proteins involved in acetate reduction to methane included the AMP forming acetyl-CoA synthetase (ACS), assigned to *M. concilii*. This protein catalyses the first step of acetoclastic methanogenesis through initial activation of acetate to acetyl-CoA.

Interestingly, this protein was encoded by separate genes, *acs*2 (*mcon*_0558) and *acs*3 (*mcon*_0559). While *acs*2 was upregulated in low-temperature biomass (R1/R2 ratio: 0.38; R1/R3 ratio: 0.22), *acs*3 (mcon_0559) was more prominent in mesophilic biomass (R1/R2 ratio: 15.8; R1/R3 ratio: 8.5) (Table 2.4). Differential expression of *acs* genes has been documented previously by Zhu *et al*. (2012), where it was reported that *acs*2 and *acs*3 transcripts were about nine fold higher than those of *acs*1 when *M*. *harundinacea* was grown in mesophilic conditions with acetate as sole carbon source. The protein expression profile of ACS2 and ACS3 found in the present study suggests that environmental factors, such as temperature, may influence expression profiles of homologous genes in *Methanosaeta* species. Indeed, differential expression of ACS may provide *Methanosaeta* with an adaptation capacity in response to environmental perturbations as proposed by Smith & Ingrim-Smith. (2007). There was a lack of acetate metabolism proteins associated with *Methanosarcina*. This could potentially be explained by the higher affinity of *Methanosaeta* for acetate in low concentrations compared to *Methanosarcina* (Jetten *et al*., 1990). The primary step in acetate activation is not conserved between the two acetoclastic methanogenic genera, with *Methanosarcina* using an acetate kinase/phosphotransacetylase (AK/PTA) system (Berger *et al*., 2012). Thus, it may be possible that differential ACS expression may also be a factor in the prevalence of *Methanosaeta*, but further analysis would be required to determine significance e.g. quantitative real-timer PCR targeting *acs* genes from *Methanosaeta* in low-temperature context. The ASC reaction is driven by the hydrolysis of pyrophosphate (PPi) by inorganic pyrophophatase. A manganese-dependent inorganic pyrophosphatase, assigned to *M*. *concilii* was detected in this study, and its level of expression was conserved for each biomass sample (Table 2.4). In recent studies investigating possible energy conservation strategies of *Methanosaeta*, it has been suggested that the energy dissipated through hydrolytic cleavage of PPi may be used for possible proton translocation (Smith & Ingrim-Smith, 2007), or phosphorylation of cellular compounds (Berger *et al*., 2012). A second protein essential in acetate metabolism is CO dehydrogenase/acetyl-CoA synthase (CODH/ACS), which oxidises the carbonyl group of acetyl-CoA to CO_2 with the methyl group transferred to the cofactor tetrahydrosarcinapterin (H_4SPT) with subsequent methane formation (Fig. 2.4). Three subunits of this protein (α, β & γ) were all assigned to *M*. *concilii*

and were conserved for each bioreactor sample (Table 2.4), thus highlighting the important functional role *Methanosaeta* has in acetate metabolism in LTAD systems.

Interestingly, a bifunctional protein formaldehyde activating enzyme/3-hexulose-6-phophate synthase (Fae/Hps) was also identified and assigned to *M. concilii*, which was upregulated in low-temperature biomass (R1/R2 ratio: 0.39; R1/R3 ratio: 0.38; Table 2.4). The Fae enzyme catalyses the condensation of formaldehyde and H_4MPT to methylene-H_4MPT and has been widely studied in relation to methylotrophic bacterial metabolism of methanol and other C_1-compounds where formaldehyde is formed as an intermediate of the pathway (Chistoserdova *et al*., 1998; Groenrich *et al*., 2005). The HPS protein has been characterised in the ribulose monophosphate (RuMP) pathway, where it catalyses the Mg^{2+}-dependent aldol condensation between formaldehyde and ribulose 5-phophate to form D-aribino-3-hexulose 6-phophate in methylotrophic bacteria (Kato *et al*., 2006). However, in some archaea it has been suggested that reverse 'RuMP pathway' may be employed for pentose phosphate synthesis with the formation of ribulose 5-phosphate (Goenrich *et al*., 2005, Soderberg, 2005). This reverse RuMP pathway would result in formaldehyde production, which may be detoxified by the Fae protein and used in subsequent methane formation via the CO_2 reduction pathway (Fig. 2.4). There was also a glycolytic enolase protein identified in this dataset, assigned once more to *M. concilii* and conserved for each temperature (Table 2.4, Fig. 2.4). It is possible that fructose-6-phosphate may have been produced from RuMP pathway in the forward direction (HPS protein identified separate from Fae complex) or by ribulose-5-phophate conversions via the pentose phosphate pathway. Thus, subsequent fructose-6-phosphate conversion via glycolysis may be occurring in these samples with resulting pyruvate formation (Fig. 2.4). This correlates with the finding that the expression of 3-isopropylmalate dehydratase, which functions in L-leucine biosynthesis as part of pyruvate metabolism, was conserved for each temperature (Table 2.4, Fig. 2.4,).

It has, moreover, been revealed that *Methanosaeta* species also hold the capacity for CO_2 reduction to methane (Smith & Ingrim-Smith, 2007). Zhu *et al*. (2012), reported

that 1% of methyl carbon of acetate was oxidised to CO_2 by *M. harandiacea* in C_{13} labelling experiments. It was suggested that reducing equivalents such as reduced F_{420} may be used for important cellular processes (e.g. cell biomass biosynthesis) by availing of this change in oxidative carbon flux. Based on these reports, and the findings of the present study, it is emerging that this acetoclastic methanogen is more metabolically diverse than previously understood. Further directed studies should be undertaken to reveal the full extent of its genetic flexibility and the potential impact of this information on its biotechnological exploitation.

The prevalence of *Methanosaeta* in this metaproteomic data set was mirrored by phylogenetic analysis of 16S rRNA gene where it was seen, for example, all samples had $>10^9$ copies $gVSS^{-1}$ (Fig. 2.3). Eight OTUs were closely aligned to *Methanosaeta* species in the clone library data, increasing in community composition (%) in synchronisation with temperature decrease (Table 2.3).

Methanogenesis from CO_2 was also found to be an important metabolic process in all biomass samples as indicated by the presence of coenzyme F_{420} dependent N5, N10-methenyl-H_4MSPT reductase. This protein was assigned to a *Methanobacterium* species and was significantly more abundant in R1 biomass compared to the two low-temperature biomass samples (R1/R2 ratio: 2.9; R1/R3 ratio: 2.3; Table 2.4). Clone library data supported this conclusion with an OTU closely aligned to *Methanobacterium beijengense* comprising 30% of the archaeal community in R1 biomass (Table 2.3, Fig. 2.2). The *Methanobacteriales* order as with *Methanosaetaceae* was detected above 10^9 copies $gVSS^{-1}$ in all biomass samples (Fig. 2.3). The *Methanobacteriales* order was also represented in the metaproteomic data set by numerous species found to express the protein methyl-coenzyme M reductase (MCR; Table 2.4). This protein catalyses the reduction of methyl-coenzyme M and coenzyme B to methane and the corresponding heterosulphide, representing the final step in methanogenesis. Varying protein expression ratios related to these species suggest a possible functional shift in the representatives of the *Methanobacteriales* order in relation to temperature, e.g. the *M. thermautotrophicus* MCR protein was significantly expressed in R2 biomass

sample (R1/R2 ratio; 0.45, R2/R3 ratio; 2.6) with the reverse recorded for the MCR assigned to M. marburgensis (R1/R2 ratio; 43.9, R2/R3 ratio; 0.07). The finding that *M. thermautotrophicus* was functioning significantly at 15°C is surprising, considering its growth range is characterized as between 30 to 70°C (Smith & Ingrim-Smith, 1997).

Molecular chaperones and regulation proteins were prominent in the metaproteomic dataset (Table 2.4). These included a GroEL chaperone assigned to the *α-proteobacteria* species *Geobacter bemidjiensis* and a DnaK chaperone assigned to *M. concilii*, which were both expressed at a higher level in mesophilic biomass (Table 2.4). In a study on *Methanococoides burtonii,* molecular chaperones were found to aid in protein folding and maintaining integrity of protein structure during heat induced stress, while protein degradation through proteolysis is preferred during cold stress conditions (Williams *et al*., 2011). A proteasome subunit was evident in the dataset at elevated levels in R2 when compared with R1 biomass sample (R1/R2 ratio: 0.40; Table 2.4). Interestingly, this protein was assigned to *Methanosarcina barkeri*. The *Methanosarcina* genera were not detected in any archaeal clone libraries (Table 2.3, Fig. 2.2) and were below the level of quantification for the qPCR assay (Fig. 2.3). The detection of a protein assigned to this acetoclastic methanogen suggests a possible functional role for it in LTAD, highlighting the benefit of integrating additional functional approaches in conjunction with DNA characterisation methods. In general, it is conceivable that microbial groups present in low numbers may have functional importance in LTAD and in many other environmental settings. Although the extent of the contribution of *Methanosarcina* to methanogenesis in the bioreactors was not possible to infer from this study, a future focus of research should be to relate findings of potential functional significance identified by metaproteomics to actual functions in the biomass, through methods such as stable isotope probing (Leuders *et al*., 2004).

A host of cell envelope proteins were also detected, which included a periplasmic solute-binding protein assigned to the *Actinobacteria* species *Segniliparus rotundus*, which was significantly expressed at a higher level in mesophilic biomass compared

to low-temperature biomass (R1/R2 ratio; 4.1, R1/R3 ratio; 3.9; Table 2.4). The presence of an S-layer related duplication protein essential for cell signalling and surface interactions (Beveridge *et al*., 1997) assigned to *M*. *concilli* and conserved for each temperature, reaffirms this groups ability to adapt in sub-mesophilic conditions.

2.5 Conclusions

In conclusion, a metaproteomic approach uncovered key insights into metabolic pathways undertaken in LTAD bioreactors. Evidence was shown that *Methanosaeta* are a key functional group in AD bioreactors operated at low-temperatures treating a synthetic VFA wastewater. The possibility that important functional groups may be under-represented through standard DNA-based methods is suggested e.g. *Methanosarcina* present in metaproteomics, absent from clone libraries and below the level of quantification in qPCR analysis. Furthermore, detection of organisms using PCR-based methods may not necessarily equate to functional importance in LTAD e.g. a *Methanocorpusculum*-like clone comprised 19% of R2 archaeal clone library yet were not represented in metaproteomic data. The lack of *Methanocorpusculum* protein detection in sub-mesophilic biomass is intriguing due to numerous studies reporting on increased numbers associated with a temperature drop in LTAD studies (McKeown *et al*, 2009; O'Reilly *et al*, 2009).Thus, requires further analysis to uncover the reasons behind this finding e.g. enrichment of psychrophilic representatives from LTAD biomass for functional characterisation. The low number of proteins assigned to bacterial species was also evident and may require further examination- for example through a linked DNA and RNA phylogenetic approach to uncover active component of bacterial consortium. Future studies in LTAD should also consider the integration of a matched metagenomic dataset, which would add significant support to the clarification of metaproteomic data. However, due to time constraints and the cost of metagenome assembly, it may be more beneficial to utilise a closely related metagenome in an iterative approach as undertaken by Roojers *et al*, (2011). Two AD metagenomic datasets have recently been generated from an agricultural biogas fermenter, which could also be used in this approach. (Jaenicke *et al*, 2011)

2.6 References

Abram, F., Gunnigle. E. and O'Flaherty, V. (2009) Optimisation of protein extraction and 2-DE for metaproteomics of microbial communities from anaerobic wastewater treatment biofilms. Electro. **30**(23), 4149-4151.

Abram, F., Enright, A.M., O'Reilly, J., Botting, C.H., Collins, G. and O'Flaherty, V. (2011) A metaproteomic approach gives functional insights into anaerobic digestion. *J*. *Appl*. *Microbiol*. **110**, 1550-1560.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic*. *Acid*. *Res*. **25**, 3389-3402.

Berger, S., Welte, C. and Deppenmeier, U. (2012) Acetate activation in Methanosaeta thermophile: Characterisation of the key enzymes pyrophosphate and acetyl-CoA synthetase. Hindawi Pub. Corp. *Archaea*. 315153.

Beveridge, T. J., Pouwels, P.H., Sára, M., Kotiranta, A., Lounatmaa, K., Kari, K., Kerosuo, E., Haapasalo, M., Egelseer, E.M., Schocher, I., Sleytr, U.B., Morelli, L., Callegari, M.L., Nomellini, J.F., Bingle, W.H., Smit, J., Leibovitz, E., Lemaire, M., Miras, I., Salamitou, S., Béguin, P., Ohayon, H., Gounon, P., Matuschek, M. and Koval, S.F. (1997) Functions of S-layers. *FEMS Microbiol*. *Rev*. **20**, 99-149.

Chistoserdova, L., Vorholt, J., Thauer, R.K. and Lidstrom, M.E. (1998) C1 transfer enzymes and coenzymes linking methylotrophic bacteria and methanogenic archaea. *Sci*. **281**, 99-102.

Cui, Y., Jiang, T., Guo, Z., Zhang, L. and Zhao, L. (2007) Development of anaerobic baffled reactor for domestic wastewater acidification under low-temperature. *Progr*. *in Environ*. *Science and Technol*. I, Science Press, Beijing. 810-814.

Collins, G., Woods, A., McHugh, S., Carton, M.W. and O'Flaherty, V. (2003) Microbial community structure and methanogenic activity during start-up of psychrophilic anaerobic digesters treating synthetic industrial wastewaters. *FEMS Microbiol*. *Ecol*. **46**, 159-170.

Collins, G., Foy, C., McHugh, S., Mahony, T. and O'Flaherty, V. (2005) Anaerobic biological treatment of phenolic wastewater at 15-18°C. *Water Res*. **39**, 1614-1620.

Collins, G., Enright, A.M., Scully, C., Mahony, T. and O'Flaherty, V. (2006) Application and biomolecular monitoring of psychrophilic anaerobic digestion. *Water Sci*. *Technol*. **54**, 41-47.

Connaughton, S., Collins, G. and O'Flaherty, V. (2006) Psychrophilic and mesophilic anaerobic digestion of brewery effluent: a comparative study. *Water Res*. **40**, 2503 2510.

DeLong, E.F. (1992) Archaea in coastal marine sediments. *Proc*. *Natl*. *Acad*. *Sci*. *U*.*S*.*A*. **89**, 5685-5689.

Enright, A.M., McHugh, S., Collins, G. and O'Flaherty, V. (2005) Low-temperature anaerobic biological treatment of solvent-containing pharmaceutical wastewater. *Water Res*. **39**, 4587-4596.

Enright, A.M., Collins, G. and O'Flaherty, V. (2007) Temporal microbial diversity changes in solvent-degrading anaerobic granular sludge from low-temperature (15°C) wastewater treatment bioreactors. *System*. *App*. *Microbiol*. **30**, 471-482.

Goenrich, M., Thauer, R.K., Yurimoto, H. and Kato, N. (2005) Formaldehyde activating enzyme (Fae) and hexulose-6-phosphate synthase (HPS) in *Methanosarcina barkeri*: a possible function in ribose-5-phophate biosynthesis. *Arch*. *Microbiol*. **184**, 41-48.

Gygi, S.P., Corthals, G.L., Zhang, Y., Rochon, Y., Aebersold, R., 2000. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. Proc. Natl. Sci. U.S.A. 97, 390-9395.

Jaenicke, S., Ander, C., Bekel, T., Bisdorf, R., Droge, M., Gartemann, K.H., Junemann, S., Kaiser, O., Krause, L., Tille, F., Zakrzewski, M., Puhler, A. and Schluter, A. (2011) Comparative and joint analysis of two metagenomic datasets from a biogas fermenter obtained by 454-pyrosequencing. *PLoS ONE*. **6**(1), e14519.

Jetten, M.S.M., Stams, A.J.M. and Zehnder, A.J.B. (1990) Acetate threshold values and acetate activating enzymes in methanogenic bacteria. *FEMS*. *Microbiol*. *Ecol*. **73**, 339-344.

Kato, N., Yurimoto, H. and Thauer, R.K. (2006) The physiological role of the ribulose monophosphate pathway in bacteria and archaea. *Biosci*. *Biotechnol*. *Biochem*. **70**(1), 10-21.

Lane, D.J., Pace, B., Olsen, G.J., Stahl, D.A., Sogin, M.L. and Pace, N.R. (1985) Rapid determination of 16S ribosomal RNA sequences for phylogenetic analysis. *Proc*. *Natl*. *Acad*. *Sci*. *U*.*S*.*A*. **82**, 6955-6959.

Lee, C., Kim, J., Hwang, K., O'Flaherty, V. and Hwang, S. (2009) Quantitative analysis of methanogenic community dynamics in three anaerobic batch digesters treating different wastewaters. *Water Res*. **43**, 157-165.

Leuders, T., Manefield, M. and Friedrich, M.W. (2004) Enhanced sensitivity of DNA- and rRNA-based stable isotope probing by fractionation and quantitative analysis of isopycnic centrifugation gradients. *Environ*. *Microbiol*. **6**(1), 73-78.

Liu, Y., Xu, H., Show, K. and Tay, J. (2002) Anaerobic granulation technology of wastewater treatment. *Microbiol*. *Biotechnol*. **18**, 99-113.

Madden, P., Chinalia, F.A., Enright, A.M., Collins, G. and O'Flaherty, V. (2010) Perturbation-independent community development in low-temperature anaerobic biological wastewater treatment bioreactors. *Biotechnol*. *Bioeng*. **105**(1), 79-87.

Martinez-Sosa, D., Helmreich, B., Netter, T., Paris, S. and Bischof, H.H. (2011) Anaerobic submerged membrane bioreactor (AnSMBR) for municipal wastewater treatment under mesophilic and psychrophilic temperature conditions. *Biores*. *Technol*. **102**, 10377-10385.

McHugh, S., Carton, M., Collins, G. and O'Flaherty, V. (2004) Reactor performance and microbial community dynamics during anaerobic biological treatment of wastewater at 16-37°C. *FEMS Microbiol*. *Ecol*. **48**, 369-378.

McKeown, R., Scully, C., Enright, A.M., Chinalia, F.A., Lee, C., Mahony, T., Collins, G. and O'Flaherty, V. (2009) Psychrophilic methanogenic community development during long-term cultivation of anaerobic granular biofilms. *ISME*. *J*. **3**, 1231-1242.

O'Farrell, P.H. (1975) High resolution two-dimensional electrophoresis of proteins. *J*. *Biol*. *Chem*. **250**, 4007-4021.

O'Reilly, J., Lee, C., Collins, G., Chinalia, F.A., Mahony. T. and O'Flaherty, V. (2009) Quantitative and qualitative analysis of methanogenic communities in mesophilically and psychrophilically cultivated anaerobic granular biofilms. *Water Res*. **43**, 3365-3374.

Roojers, K., Kolmeder, C., Juste, C., Dire, J., de Breen, M., Boeren, S., Galan, P., Beauvallet, C., de Vos, W. and Schaap, P. (2011) An iterative workflow for mining the human intestinal metaproteome. *BMC*. *Genomics*. **12**, 6.

Scully, C., Collins, G. and O'Flaherty, V. (2006) Anaerobic biological treatment of phenol at 9.5-15°C in an expanded granular sludge bed (EGSB)-based bioreactor. *Water Res*. **40**, 3737-3744.

Shilova, I.V., Seymour, S.L., Patel, A.A., Loboda, A., Tang, W.H., Keating, S.P., Hunter, C.L., Nuwaysir, L.M. and Schaeffer, D.A. (2007) The paragon algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Molec*. *Cell*. *Prot*. **6**(9), 1638-1655.

Siggins, A. (2010) Low-temperature anaerobic digestion of trichloroethylene. PhD thesis. NUI Galway, Ireland.

Siggins, A., Enright, A.M. and O'Flaherty, V. (2011a) Methanogenic community development in anaerobic granular bioreactors treating trichloroethylene (TCE) contaminated wastewater at 37°C and 15°C. *Water Res*. **45**, 2452-2462.

Siggins, A., Enright, A.M. and O'Flaherty, V. (2011b) Temperature dependent (37°C-15°C) anaerobic digestion of a trichloroethylen-contaminated wastewater. *Bioresource*. *Technol*. **102**, 7645-7656.

Siggins, A., Enright, A.M. and O'Flaherty, V. (2011c) Low-temperature (7°C) anaerobic treatment of a trichloroethylene-contaminated wastewater: Microbial community development. *Water Res*. **45**, 4035, 4046.

Smith, K.S. and Ingram-Smith, C. (2007) Methanosaeta, the forgotten methanogen? *Trends Microbiol*. **15**(4), 150-155.

Soderberg, T. (2005) Biosynthesis of ribose-5-phosphate and erythrose-4-phosphate in archaea: a phylogenetic analysis of archaeal genomes. *Archaea*. **1**, 347-352.

Stackebrandt, E. and Goodfellow M. (1991) Nucleic Acid Techniques in Bacteria Systematics. Chichester: Wiley.

Stamatakis, A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinform*. **22** (21), 2688–2690.

Williams, T.J., Lauro, F.M., Ertan, H., Burg, D.W., Poijak, A., Raftery, M. and Cavicchioli. (2011) Defining the response of a microorganism to temperatures that span its complete growth temperature range (-2°C to 28°C) using multiplex quantitative proteomics. *Environ*. *Microbiol*. **13** (8), 2186-2203.

Wilmes, P. and Bond, P.L. (2004) The application of two-dimensional polyacrylamide gel electrophoresis and downstream analysis to a mixed community of prokaryotic microorganisms. *Environ*. *Microbiol*. **6**, 911-920.

Yu, Y, Lee, C., Kim, J. and Hwang, S. (2005) Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnol*. *Bioeng*. **89**, 670-679.

Zhang, D., Zhu, W., Tang, C., Suo, Y., Gao, L., Yuan, X., Wang, X. and Cui, Z. (2012) Bioreactor performance and methanogenic population dynamics in a low-temperature (5-18°C) anaerorobic fixed-bed reactor. *Biores*. *Technol*. **104**, 136-143.

Zhu, J., Zheng, H., Guomin, A., Zhang, G., Liu, D., Xiaoli, L. and Dong, X. (2012) The genome characteristics and predicted function of methyl-group oxidation pathway in the obligate acetoclastic methanogens, *Methanosaeta* spp. *PLoS ONE*. **7** (5), e367.

Chapter 3

A functional approach to uncover the low-temperature adaptation strategies of the archaeon *Methanosarcina barkeri*

A condensed version of this chapter was submitted to Applied Environmental Microbiology:

Gunnigle, E., McCay, P., Fuszard, M., Botting, C.H., Abram, F., O'Flaherty, V. (2012). A functional approach to uncover the low-temperature adaptation strategies of the archaeon *Methanosarcina barkeri*. Submitted Dec 2012.

Abstract

Low-temperature anaerobic digestion (LTAD) technology is underpinned by a diverse microbial community. The methanogenic archaea represent a key functional group in these consortia, being responsible for CO_2 reduction as well as acetate and methylated C_1 metabolism with subsequent biogas (40-60% CH_4, 30-50% CO_2) formation. However, the cold adaptation strategies, which allow methanogens to function efficiently in LTAD remains unclear. In this study, a pure culture proteomic (iTRAQ, 2-DGE) approach was employed to uncover the sub-mesophilic functional characteristics of a *Methanosarcina barkeri* (optimum growth temperature, 37˚C), which has been detected in LTAD bioreactors. Two experimental approaches were used to gain some insight into the low-temperature adaptation of *M*. *barkeri*. The first experimental setup aimed to characterise a low-temperature shock response (LTSR) of *M*. *barkeri* DSMZ 800^T grown initially at 37°C with a temperature drop to 15°C, while the second experimental approach aimed to examine the low-temperature adaptation strategies (LTAS) of the same strain, when grown solely at 15°C. This latter experiment employed cell viability and growth measurements (OD_{600}), which directly compared *M*. *barkeri* cells grown at 15°C with those grown at 37°C. During the LTSR experiment, a total of 127 proteins were detected in 37°C and 15°C samples with 20 proteins differentially expressed with respect to temperature, while in the LTAS experiment 39% of proteins identified were differentially expressed between phases of growth. Functional categories included methanogenesis, cellular information processing and chaperones. By applying a polyphasic approach (proteomics and growth studies), insights into the low-temperature adaptation capacity of this mesophilically characterised methanogen were obtained, which suggest that the metabolically diverse *Methanosarcinaceae* could be functionally relevant for LTAD systems.

3.1 Introduction

Archaea are ubiquitous in low-temperature habitats such as polar marine waters (Church *et al*., 2003; Galand *et al*., 2006), alpine lakes (Pernthaler *et al*., 1998), permafrost (Wagner *et al*., 2003) and glacier ice (Battin *et al*., 2001). Methanogenic archaea represent the most characterised psychrophilic archaeal group (Cavicchioli, 2006). As such, low-temperature methanogenesis has been the focus of many studies, such as those focusing on measuring methanogenic contribution to the global warming of cold areas (Hoj *et al*., 2008), and also from an astrobiological viewpoint, where the ability of methanogens to survive in cold anoxic conditions has made them candidates as Earth analogues for extra-terrestrial life (Morozova & Wagner, 2007).

Low-temperature methanogenic activity is also important from a biotechnological viewpoint, such as its application of low-temperature anaerobic digestion (LTAD; Collins *et al*., 2006). Evidence of efficient LTAD treatment of wastewaters has been recorded in laboratory-scale trials, which directly compared low-temperature bioreactor performance (chemical oxygen demand removal [COD], biogas production) with traditional mesophilic configurations (Connaughton *et al*., 2006; O'Reilley *et al*., 2009). Experiments, which comprised an initial mesophilic (37˚C) bioreactor operation phase, followed by a decrease to low-temperature conditions (≤15˚C) have also been undertaken (McHugh *et al*., 2004; McHugh *et al*., 2006). In these studies, low-temperature bioreactors achieved comparable performance levels to mesophilic systems, after an initial period of adaptation. As a mesophilic inoculum was used to seed these bioreactors, a psychrotolerant capacity was deemed to be evident in the mixed microbial consortia, underpinning these bioreactors. However, there still remains a significant knowledge gap relating to low-temperature methanogenic adaptation strategies, which requires further elucidation for the optimisation of LTAD systems.

There are three primary modes of methanogenic metabolism, based on: CO_2-reduction; acetate decarboxylation; and methylotrophic activity (e.g. reduction of

methylamines). Acetoclastic (acetate decarboxylation) methanogenesis has been recorded as being the primary methanogenic pathway in low-temperature environments, including bioengineered systems (McKeown *et al*., 2009). The order *Methanosarcinales* includes the only two known acetoclastic families, the *Methanosaetaceae* and the *Methanosarcinaceae*. The former has been historically categorised as a strict acetoclastic methanogen, although a recent genomic study highlighted the metabolic capacity for possible methyl group oxidation in three sequenced *Methanosaetaceae* strains (Zhu *et al*., 2012). Nevertheless, this group has a minimum threshold concentration of ~1mM acetate and has been documented to outcompete *Methanosarcinaceae* in environments where acetate concentrations are low (Fey *et al*., 2000; Griffin *et al*., 1998). In addition to acetate, the *Methanosarcinaceae* have the ability to utilise methylated compounds such as methanol and methylamines, with some species also able to use H_2/CO_2 as a carbon and energy source.

In previous LTAD studies, *Methanosaeta*-like organisms have been found to be the most abundant acetoclastic methanogens, and their presence has been correlated with both granular sludge integrity and high levels of process efficiency (Diaz *et al*., 2006; Liu *et al*., 2002). By contrast, LTAD methanogenic communities are usually characterised by low levels of *Methanosarcinaceae*. This was demonstrated in an LTAD study (Siggins *et al*., 2011), where *Methanosarcinaceae* was detected below the quantification limit of the 16S rRNA gene assay, whereas the *Methanosaetacaeae* comprised 75% of total measured methanogenic 16S rRNA gene concentrations. However, a marked increase in *Methanosarcinaceae* has been recorded in previous LTAD studies, particularly during periods of acetate accumulation, and the organism may thus play an important role in ensuring process stability during transient perturbations (McKeown *et al*., 2009; O'Reilley *et al*., 2009). Moreover, in a recent proteomic investigation on LTAD bioreactor samples, proteins assigned to *Methanosarcinaceae* were identified, although no corresponding representatives were noted in 16S rRNA gene clone libraries (Abram *et al*., 2011). This finding suggested that *Methanosarsinaceae* may play a functional role in LTAD, even if cell numbers are relatively low by comparison to *Methanosaetaceae* and other methanogenic populations. Interestingly, no proteins related to acetate

utilisation in *Methanosarsinaceae* were recovered during this trial (Abram *et al*., 2011), suggesting that the utilisation of substrates, other than acetate, may have been the basis for the persistence and growth of the organism during LTAD.

The *Methanosarcinaceae* species *Methanosarcina barkeri* is a metabolically versatile methanogen, which can grow on H_2/CO_2, methanol, various methylamines and acetate as carbon and energy sources (Bock & Schonheit, 1995). This species has previously been the focus of a variety of studies, including growth experiments (Mazumder *et al*., 1986; Muller *et al*., 1986), genomic investigations (Feist *et al*., 2006; Maeder *et al*., 2006), and enzymatic characterisation (Buchenau *et al*., 2004; Goenrich *et al*., 2005). In the present study, we employed a systematic, proteomics-based, approach to characterise the low-temperature adaptive strategies of a methanol- and H_2/CO_2-metabolising *M. barkeri*, strain DSMZ 800^T, which was isolated from an anaerobic digester. Our hypothesis was that the organism would display an adaptive capacity to allow growth and high rates of utilisation of methanol and H_2/CO_2 under low-temperature conditions.

3.2 Materials and Methods

3.2.1 Strain information, media and inoculum preparation

The archaeal strain *Methanosarcina barkeri* DSMZ 800^T (DSMZ, Braunschweig, Germany) was used throughout this study. Inoculum preparation was achieved using medium DMS 120 (pH 6.8; Nomura *et al*., 2008). After 40 ml of medium was added, each 60 ml vial was sealed with a butyl rubber bung and flushed with N_2/CO_2 (80:20 v/v) gas. Filter sterilised reducing agents L-cysteine (2.5 mM) and Na_2S x $9H_2O$ (1.2 mM) were added after autoclaving. In order to investigate substrate influence on growth, selected vials were supplemented with either methanol (156 mM, 50% v/v stock, filter sterilised), H_2/CO_2 (80:20 v/v) or methanol + H_2/CO_2. Vials were inoculated with 2.5% (v/v) stock culture and incubated at 37°C on a shaker.

3.2.2 Batch culture setup

When cells reached stationary phase (OD_{600} of ~ 2), inoculum was added anaerobically to fresh vials to an OD_{600} of ~ 0.05 in 40 ml of media. Separate vials were then incubated at 15°C and 37°C for comparative analysis on a shaker. OD_{600} measurements were taken at time intervals for each temperature, while separate vials for headspace methane measurements (%; APHA, 1998) were run in parallel. This experiment aimed to give direct insights into the capacity of *M. barkeri* to adapt to low-temperatures, and therefore will be referred to as 'low-temperature adaptation strategies' (LTAS) from this point on. Proteomic analysis was also carried out on a subset of samples (Table 3.1) and, for this purpose cells from 60 ml of culture were harvested by first flash freezing with liquid nitrogen before storing at -80°C. Samples were then thawed and centrifuged at 8,000 *g* for 10 mins at 4°C. The supernatant was discarded and the recovered biomass pellet (approximately 1.5g wet weight per sample) was stored at -80°C until use.

3.2.3 Continuous culture setup

Continuous cultures of *M. barkeri* were established in a modified 500 ml Schott bottle (fitted with a 3-port lid, Fisher Scientific) with a working volume of 400±100 ml, connected to both a medium reservoir and a waste reservoir. Anaerobic conditions were maintained in the vessel by sealing with butyl stoppers, using gas impermeable tubing connected to medium reservoir and by regularly flushing with N_2/CO_2 (80:20 v/v) gas. The sterile medium was allowed to acclimatise at 37°C overnight without shaking prior to inoculation with 2.5% (v/v) ml of culture inoculum. The flow-rate from the reservoir was controlled by a peristaltic pump (Watson-Marlow 205U, USA) and was monitored and adjusted to maintain a steady dilution rate of 2.0 h^{-1}. The continuous culture system was operated at 37°C, allowing a continuous culture to establish (3.5 days, OD_{600} ~ 1.9), whereafter the temperature was dropped to 15°C. The system was further operated for 3 days. Samples for protein extraction (60 ml) were taken immediately before the temperature drop event and also at the end of 15°C operation (Table 3.1). This experiment aimed to uncover the functional response of *M. barkeri* directly after a

temperature drop to 15°C, therefore will be referred to as 'low-temperature shock response' (LTSR) from this point on.

3.2.4 Cell viability measurements

The viability of cells during LTAS experiment was investigated using a LIVE/DEAD® *Bac*Light™ kit (Invitrogen; [Hao *et al*., 2009]; Table 3.1). Anaerobic conditions were maintained in samples during incubation with *Bac*Light™ probes by sealing with N_2/CO_2 (80:20, v/v) gas. A Nikon Eclipse E600 epifluorescent microscope (Carl Zeiss, Oberkochen, Germany) was used to assess cell viability. Live cells were indicated by green fluorescence (BS2 filter; excitation 450-490 nm, emission LP515 nm) while dead cells were indicated by red fluorescence (CY3 filter; excitation 546±12 nm, emission at LP590 nm). Ten images of microscope field (100x) were taken for each sample, and visualised using QCapture Pro software.

3.2.5 Protein extraction and quantification

M. *barkeri* proteins were extracted from samples using a sonication method adapted from a previous LTAD proteomic study (Abram *et al*., 2011; Table 3.1). Briefly, cell pellets were resuspended with 5 ml non-interfering triethylammonium bicarbonate (TEAB) buffer. The cells were then disrupted by sonication (MSE soniprep 150) at 40% amplitude for 30 seconds on ice. Fifteen pulses were applied with 30 second intervals to prevent thermal damage of cellular proteins. In order to remove any protein precipitates and cellular debris, lysates underwent centrifugation at 10,000 *g* for 30 mins at 4°C with the resulting pellets discarded. Protein precipitation from the supernatants was achieved through incubation with ice cold acetone (3:1 v/v) at -20°C for 1 hour after which the samples underwent centrifugation at 10,000 *g* for 15 mins. The supernatants were discarded and pellets resuspended with 200 µl TEAB buffer. Protein quantification was undertaken using a Calbiochem Non-Interfering Protein Assay™ kit (Merck KGaA, Darmstadt, Germany) following the manufacturer's instructions.

3.2.6 iTRAQ labelling

For each study (LTAS & LTSR), replicate protein extracts from specified time-points were used in an 8 plex-iTRAQ (isobaric Tags for Relative and Absolute Quantification) labelling experiment (Table 3.1). One hundred µg protein extract, in 20 µl samples, were labelled with iTRAQ reagents, according to manufacturer's guidelines (ABSciex, Foster City, CA). Briefly, each sample was reduced, denatured and cysteine residues were blocked before digesting with 10 µl of 25 µg/µl trypsin solution (1:20; Promega, Madison, WA) overnight, at 37°C. The digested peptides were then labelled with iTRAQ tags as follows: i) in the LTAS experiment: duplicate samples (I) - 113 and 114 tags, duplicate samples (II) - 115 and 116 tags, duplicate samples (III) -117 and 118 tags, and duplicate samples (IV) -119 and 121 tags; ii) in LTSR experiment only four iTRAQ tags were required; duplicate 37°C samples - 113 and 114 tags, duplicate 15°C samples - 115 and 116 tags. For each experimental setup the samples were combined and subjected to peptide separation by cation exchange chromatography followed by LC-MS/MS analysis.

3.2.7 Cation exchange peptide separation and LC-MS/MS analysis

After trypsin digestion and labelling, peptides from individual samples were combined and resuspended in 1 ml cation exchange load buffer (10 mM KH_2PO_4 pH 3.0 in 25 % acetonitrile). The peptides were then separated by strong cation exchange chromatography on a PolySulfoethyl A column (Poly-LC) with an increasing salt gradient from 0 – 0.5 M over 30 min with fractions collected every 30 sec. Fractions were pooled to give 9 fractions of approximately equal peptide concentration, as judged by inspection of the chromatogram, evaporated to dryness, resuspended in 0.1% trifluoroacetic acid and desalted (PepClean C18 spin columns, Thermo Scientific). Each fraction was further separated through reverse phase chromatography before mass spectrometric analysis. This was achieved using an Eksigent nanoLC-Ultra coupled to a nanoflex cHiPLC (ABSciex, Foster City, CA) equipped with a 200µm x 0.5mm ChromXP C18-CL 3µm 120Å trap and 75µm x 15 cm ChromXP C18-CL 3µm 120Å column, using a gradient of increasing acetonitrile concentration, containing 0.1 % formic acid (15-40% acetonitrile in 40 min, 40-95%

in a further 10 min, followed by 95% acetonitrile to clean the column). The eluent was sprayed into an ABSciex TripleTOF 5600 mass spectrometer (ABSciex, Foster City, CA) and analysed in Information Dependent Acquisition (IDA) mode, performing 0.25 sec of MS followed by 2 sec MSMS analyses of the 20 most intense peaks seen by MS. These masses were then excluded from analysis for the next 13 sec. A rolling collision energy adjusted to 10 units higher than that normally used for peptides was employed to provide sufficient peptide fragmentation and generation of the iTRAQ reporter groups.

3.2.8 Two-dimensional gel electrophoresis (2-DGE)

Eight gels were run in total corresponding to two duplicate independent extractions and two technical replicates from 37°C and 15°C LTAS samples following the protocol described in a previous study (Abram *et al*., 2011; Table 3.1). Gel images were processed and analysed with PDQuest-Advanced software, version 8.0.1 (BioRad). Spot counts were obtained using the spot detection wizard enabling the Gaussian model option as recommended by the manufacturer. Ratios of spot intensities were determined for each sample. Protein expression ratios greater than 1.5-fold were considered significant. Proteins of interest were excised from the gels and subjected to in-gel digestion prior to analysis using nanoflow liquid chromatography-electrospray ionization tandem mass spectrometry (nLC-ESI-MS/MS: Abram *et al*., 2011).

3.2.9 Protein identification and quantification

The MS/MS data recovered from iTRAQ and 2-DGE LC-MS/MS were analysed as follows: MS/MS data for +1 to +5 charged precursor ions, which exceeded 150 cps, were processed using the Paragon™ search algorithm (Shilova *et al*., 2007) within ProteinPilot 4.1 software (ABSciex, Foster City, CA) against an internal database comprising the *M.barkeri* Fusaro genome (DSM 804, 3616 proteins UniProt KB) along with potential contaminants (Uniprot/Swissprot, accessed 19/06/2012). The data were searched with a Detected Protein Threshold (Unused ProtScore) of >0.05,

trypsin as the cleavage enzyme, one missed cleavage, iTRAQ modification of lysine's and methyl methanethiosulfonate (MMTS) modification of cysteine's as fixed modifications and methionine oxidation selected as a variable modification. Peptide and protein lists were then exported from ProteinPilot to ProteinPilot™ Descriptive Statistics Template (PDST) (ABSciex, Foster City, CA) for data management and additional analysis (e.g. Volcano & Quant FDR worksheet). The false discovery rate (FDR) was calculated by ProteinPilot™ PDST using the embedded decoy search, and was found to be <2% for all experiments.

3.3 Results

3.3.1 Characteristics of M. barkeri grown at 37°C and 15°C

The growth of *M. barkeri* at 37°C and 15°C was compared through OD_{600} measurements with different substrate mixes (Methanol + H_2/CO_2; Methanol; H_2/CO_2; Fig. 3.1). The maximum specific growth rate (μ_{max}) was found to be dependent on both temperature and substrate (Table 3.2). The OD_{600} measurements indicated a minimal lag phase for cultures grown at 37°C, regardless of substrate, with an increase in OD_{600} correlated with an increase in CH_4% production (Fig. 3.1). After a lag phase of ~ 1 day, growth at 15°C exhibited a similar profile to that observed at 37°C. On day 3, however, a long stationary phase was recorded prior to a further steady increase in growth, from day 17. Although the μ_{max} was lower in the 15°C cultures, the H_2/CO_2 + methanol fed cultures reached similar headspace CH_4% levels in comparison to their mesophilic counterparts, peaking at 70% CH_4 (±4.3% SE; n=3) on day 41 of the incubation compared with 64% (±6.7% SE; n=3) on day 14 in 37°C vials (Fig. 3.1). The *M. barkeri* cultures with H_2/CO_2 as a sole substrate exhibited poor growth and % methane headspace values at both temperatures tested, compared with cultures fed with H_2CO_2 + methanol, or with and methanol only (Fig. 3.1).

Table 3.1. Overview of *M. barkeri* growth (H_2CO_2 + methanol as substrates) characteristics (OD_{600}, CH_4%) with and experimental approaches applied in each study

Temperature	**37°C**				**15°C**								**37°C**	**15°C**
	LTAS[a]				(I)	(II)			(III)			(IV)	**LTSR**[b]	
Time (days)	0.8	1.8	3.7	7	1.8	3.7	7.0	11.8	17.2	24.2	29.2	37.2	3.5	7.2
OD_{600}	0.5	0.56	2.1	1.3	0.05	0.3	0.25	0.23	0.21	0.24	0.35	0.76	1.9	0.8
CH_4%	1.8	19	46	43	1.2	1.9	2	3.7	6.7	15.6	25	46	58	46
Live/dead	●	●	●	●	●	●	●	●	●	●	●	●		
iTRAQ					●	●			●			●	●	●
2-DGE			●									●		

[a] LTAS refers to the 'low-temperature adaptation strategies' iTRAQ experiment, [b] LTSR refers to the 'low-temperature shock response' iTRAQ experiment. (I)-(IV) represent sampling points for iTRAQ analysis.

Table 3.2. . Maximum specific growth rates (μ_{max}) of *M. barkeri*

Substrate	**Temperature**	**μ_{max} (d^{-1})**	**% decrease in μ_{max}**	**Generation time (days)**
H_2CO_2 + Methanol	37°C	0.5 ± 0.03	0	1.4
Methanol	37°C	0.46 ± 0.06	8	1.5
H_2CO_2	37°C	0.11 ± 0.02	78	6.3
H_2CO_2 + Methanol	15°C	0.12 ± 0.07	0	6.2
Methanol	15°C	0.09 ± 0.03	50	10.7
H_2CO_2	15°C	0.03 ± 0.09	75	24.2

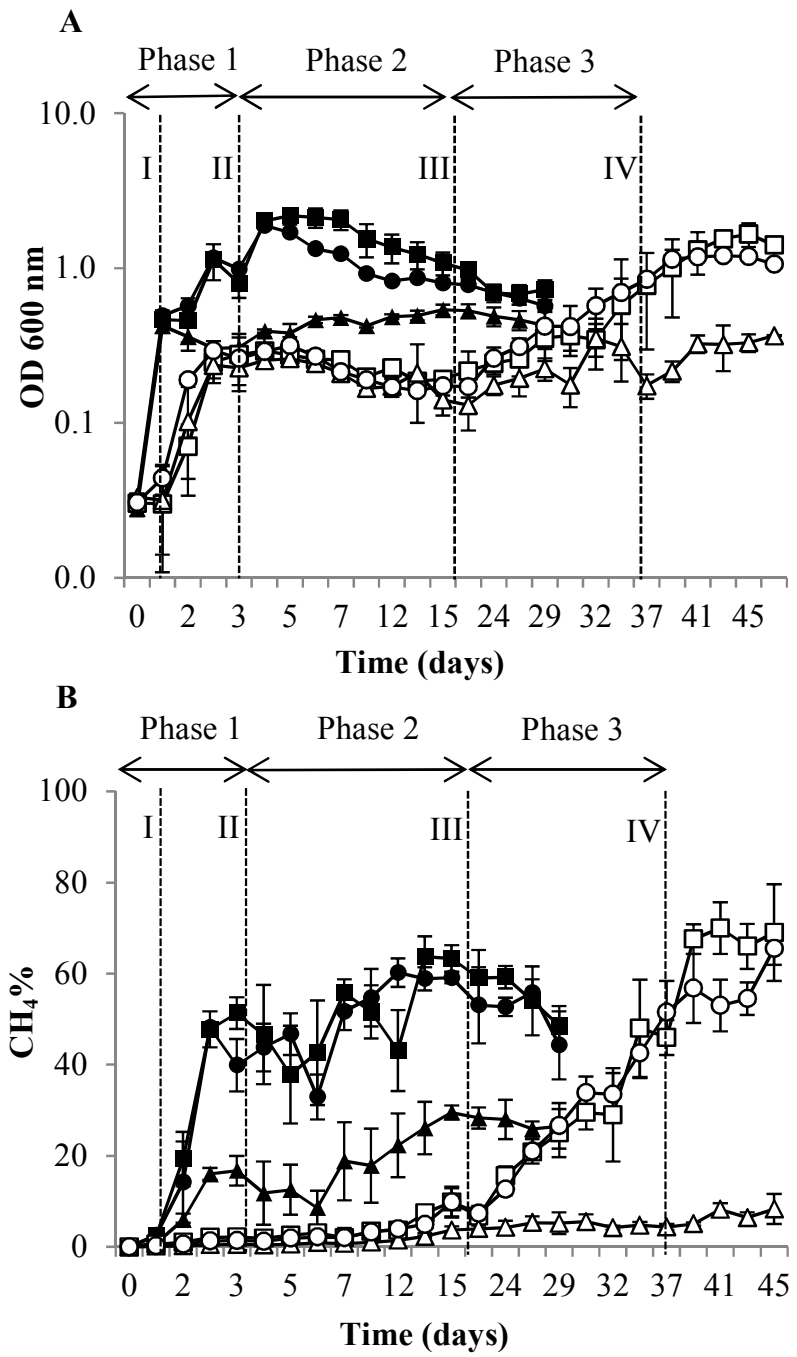


Fig. 2.1. Growth profiles of *M*. *barkeri* incubated at 37°C and 15°C. (A) Optical density (600nm) measurements and (B) CH_4% headspace measurement results in tandem with OD_{600} measurement for *M*. *barkeri* grown at 37°C with H_2CO_2 + Methanol (■), Methanol (●) and H_2CO_2 (▲) and 15°C with H_2CO_2 + Methanol (□) Methanol (○) and H_2CO_2 (△). Error bars indicate the standard deviation and are the result of at least 3 replicates. Data labels include Phase 1 (samples I & II), Phase 2 (sample III) and Phase 3 (sample IV) of 'low-temperature adaptation strategies' (LTAS) experiment.

3.3.2 Cell viability

The viable cell fractions in 37°C samples increased from 41% (±6.5% SE; n = 10) at day 0.8 to 59% (±17.8% SE; n = 10) at day 1.8 (Fig. 3.2). The total cell count peaked after 3.7 days (OD_{600} ~ 2), but the viable cell fraction did not show any significant increase with only 60% (±11.2% SE; n = 10) of total cells recorded as being viable at that point. At 15°C, however, a lower total cell count on day 7 was not reflective of the viable cell fraction, which was 79% (±13.8% SE; n = 10; Fig. 3.2). This trend continued, as recorded at 17.2 days with 84% (±9.65% SE; n = 10) of cells being identified as viable. After this observed period of minimal growth but high viability, the total cell counts increased in 15°C samples, while the viable fraction decreased to 66% (±9.5% SE; n = 10) on day 37. The spatial arrangement of *M. barkeri* aggregates from both temperatures varied with 37°C samples consisting of layered branching aggregates, while the 15°C aggregates were more consistent in size, with clusters rarely exceeding 10 µm in diameter (Fig. 3.2).

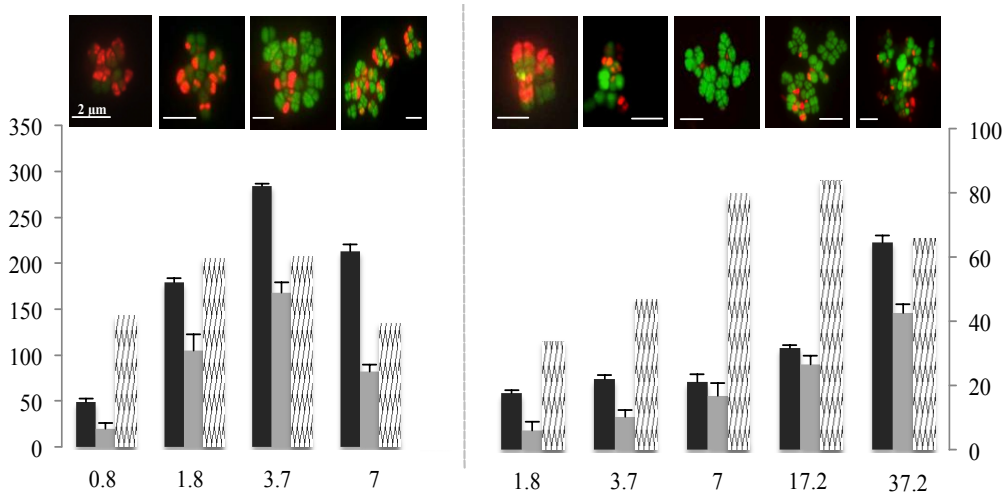


Fig. 3.2. Cell viability of *M. barkeri* grown at 37°C and 15°C. Bar chart representing total cell counts (■), viable cell counts (■) and viable fraction (%) of total cell counts (■) with representative live/dead images of *M. barkeri* taken from samples during the course of growth curve incubations. Error bars indicate the standard deviation and are the result of at least 5 replicates.

3.3.3 Proteomic analysis

In order to gain some insight into the low-temperature adaptation of *M. barkeri*, two iTRAQ experiments were performed; one involving cultures grown at 37°C and exposed to a temperature drop to 15°C (LTSR) and another experiment involving cultures grown at 15°C (LTAS). In LTSR study, a total of 127 proteins were detected in both 37°C and 15°C samples. From these, 20 proteins (16%) were found to be differentially expressed ($P \leq 0.05$) as a function of temperature, by a 1.5-fold change or more (Table 3.3). Amongst these, various functional categories were evident. Two proteins involved in CO_2 methanogenesis were detected at higher levels at 37°C when compared to 15°C (MbarA1095, MbarA1763; Table 3.3). Methanogenesis from dimethylamine and methanol showed contrasting results with a protein involved in the dimethylamine pathway (MbarA3605) identified at a higher level at 37°C sample, while the corrinoid protein involved in the methanol pathway (MbarA3637) was upregulated at 15°C. Several information processing enzymes were identified as being differentially expressed. These included an XRE domain protein (MbarA1033), which plays a key role in DNA binding (Veit *et al*., 2006), and an elongation factor 2 (EF-2) protein (MbarA1064), with both proteins upregulated at 15°C (Table 3.3). In addition, a ribosomal protein (MbarA0614) was found to be expressed at a higher level at 15°C. Also, several stress proteins were upregulated at 37°C, which included a heat shock protein (MbarA1543) and two subunits of the thermosome protein (MbarA1084, MbarA1201; Table 3.3).

In conjunction with the *M. barkeri* growth profile at 15°C, the iTRAQ samples for LTAS study were taken during the three phases observed: Phase 1 (initial adaptation [samples I & II]), Phase 2 (stationary phase [sample III]) and Phase 3 (primary growth phase [sample IV]) (Fig. 3.1, Table 3.1). In total, 104 proteins were identified in all samples with 41 (39%) found to be differentially expressed ($P \leq 0.05$, 1.5-fold difference) between two or more samples (Table 3.3).

Table 3.3. Differentially expressed proteins from iTRAQ experiments.

Gene locus	Functional description	Sequence coverage (%)	Differential abundant proteins (Ratio)					
			LTSR	LTAS				
			37°C/15°C	P1[a] vs P2 (I/III)	(II/III)	P1 vs P3 (I/IV)	(II/IV)	P2 vs P3 (III/IV)
		Methanogenesis						
		CO_2-specific						
MbarA0450	Coenzyme F_{420} hydrogenase subunit G	12	m	0.61	0.55	0.51	0.58	ns
MbarA1095	F_{420}-dependent H_4MPST dehydrogenase	34	3.27	ns	ns	0.6	0.6	0.62
MbarA1763	Formylmethanofuran dehydrogenase subunit B	13	3.19	m	m	m	m	m
		Mono/di/tri methylamine-specific						
MbarA0841	Methylcobalamin:CoM methyltransferase	8	ns	0.62	0.62	0.48	0.46	ns
MbarA3605	Dimethylamine methyltransferase	8	2.98	0.59	0.58	0.46	0.44	ns
MbarA1502	Trimethylamine methyltransferase	7	ns	0.42	0.46	0.52	0.56	1.54
		Methanol-specific						
MbarA1063	Methanol corrinoid protein	30	ns	ns	ns	0.44	0.44	0.6
MbarA3637	Methanol corrinoid protein	12	0.26	m	m	m	m	m
MbarA1064	Methanol:corrinoid methyltransferase	17	ns	ns	ns	1.62	1.58	1.62
		Common methanogenic proteins						
MbarA0893	Methyl-coenzyme M reductase subunit A	18	ns	ns	ns	0.49	0.47	0.56
MbarA1182	Methyl-coenzyme M reductase subunit B	15	ns	ns	ns	0.62	0.52	0.43
MbarA1256	Tetrahydromethanopterin S-methyltransferase subunit G	26	4.78	ns	ns	ns	ns	ns
		Cellular information processing						
		Transcription						
MbarA0507	Putative nickel-responsive regulator 1	29	2.69	ns	ns	ns	0.56	0.64
MbarA0841	DNA-directed RNA polymerase I, II and III, 7.3 kDa polypeptide	37	m	0.49	0.54	0.37	0.4	ns
		Translation						
MbarA3687	30S ribosomal protein S7P	27	m	1.65	1.56	1.52	ns	ns
MbarA1149	30S ribosomal protein S17	16	m	ns	ns	1.84	1.53	1.64
MbarA3231	50S ribosomal protein L21E	15	m	0.62	0.48	0.5	0.33	ns
MbarA0614	50S ribosomal protein L11P	16	0.29	ns	ns	0.61	0.6	0.65
MbarA3173	Aspartyl-tRNA synthetase	23	ns	1.57	1.89	ns	ns	0.45
MbarA0454	SSU ribosomal protein S19E	4	ns	0.65	0.59	0.5	0.45	ns
MbarA0623	Isocitrate dehydrogenase	4	ns	ns	ns	0.53	0.51	ns
MbarA2048	Leucyl-tRNA synthetase	17	2.58	m	m	m	m	m
MbarA3685	Elongation factor 1 subunit A	12	ns	0.67	0.56	0.62	0.6	ns
MbarA1064	Elongation factor 2	18	0.48	ns	ns	1.62	1.54	1.63
		DNA binding & repair						
MbarA1033	Transcriptional regulator, XRE family	8	0.29	ns	ns	1.52	1.51	1.54
MbarA1182	Nucleoid protein MC1	22	m	0.61	0.47	0.42	0.32	ns
MbarA1558	Archaeal histone	30	m	0.6	0.54	0.51	0.45	ns

Gene locus	Functional description	Sequence coverage (%)	Differential abundant proteins (Ratio)					
			LTSR	LTAS				
			37°C/15°C	P1 vs P2 (I/III)	(II/III)	P1 vs P3 (I/IV)	(II/IV)	P2 vs P3 (III/IV)
		General metabolism						
		Amino acid biosynthesis						
MbarA1139	3-isopropylmalate dehydrogenase	7	1.92	0.51	0.52	0.41	0.41	ns
MbarA0220	Ketol-acid reductoisomerase	9	1.78	0.54	ns	0.62	0.58	ns
MbarA1431	D-3-phosphoglycerate dehydrogenase	37	2.17	ns	ns	ns	ns	ns
MbarA1094	Phosphoserine phosphatase	18	1.84	ns	ns	ns	ns	ns
MbarA3623	Tryptophan synthase beta chain 1	12	3.19	m	m	m	m	m
		Vitamin biosynthesis						
MbarA0597	Phosphomethylpyrimidine synthase 1	10	1.66	ns	0.66	ns	ns	ns
MbarA1056	CobW protein	4	m	ns	0.54	ns	ns	ns
		Glycolysis and energy metabolism						
MbarA2189	Glyceraldehyde-3-phosphate dehydrogenase 1	6	ns	ns	ns	0.53	0.49	ns
MbarA0392	H(+)-transporting ATP synthase, subunit H	19	ns	ns	ns	0.47	0.44	0.66
		Electron transport						
Mbar A1847	Methanophenazine-reducing hydrogenase	8	ns	0.61	ns	0.42	0.46	ns
Mbar A3651	HesB protein	13	ns	0.63	0.61	0.41	0.43	0.65
		Chaperones and stress proteins						
MbarA1084	Thermosome subunit A	22	2.67	0.62	ns	0.5	0.48	0.57
MbarA1201	Thermosome subunit B	15	1.54	ns	ns	0.52	0.52	0.66
MbarA3433	Chaperone protein DnaK	8	ns	0.56	0.55	ns	ns	1.56
MbarA1543	60 kDa chaperonin	32	1.67	m	m	m	m	m
MbarA2503	Proteasome α-subunit	12	ns	0.56	0.53	0.47	0.44	ns
MbarA2249	Peptidyl-prolyl isomerase	20	ns	0.45	0.42	ns	ns	1.64
MbarA0213	Prefoldin subunit B	21	ns	1.56	ns	ns	ns	ns
MbarA0825	Universal stress protein	11	3.24	ns	ns	0.46	0.45	1.66
		Unknown						
MbarA0488	Putative uncharacterised protein	8	ns	0.52	0.64	0.5	0.6	ns
MbarA1026	Putative uncharacterised protein	29	m	ns	ns	1.57	1.51	1.67
MbarA0376	Putative uncharacterised protein	17	m	ns	ns	0.57	0.51	0.64
Mbar A1027	Putative uncharacterised protein	46	m	1.53	1.56	2.38	2	2.12
MbarA3219	Putative uncharacterised protein	15	m	ns	ns	0.41	0.4	ns

Where protein data are unavailable for a particular sample, the letter m is used to indicate missing values. Where no statistically significant differences for protein values were evident, the letters ns were used. [a] P1 refers to Phase 1 (samples I & II), P2:Phase 2 (sample III), P3:Phase 3 (sample IV).

Figure 3.3 illustrates differentially expressed protein profiles for each of the phases investigated in LTAS study. Proteins relating to CO_2 and methanol methanogenesis as well as methylamine specific proteins were found to be significantly expressed in one or more of these phases. A methanol metabolism protein (MbarA1064) was upregulated in Phase 1 and Phase 2 samples in comparison to Phase 3 sample. A trimethylamine protein (MbarA1502) was found to be expressed at a higher level in the Phase 2 sample when compared with other samples. However, the majority of methanogenesis proteins were upregulated in the Phase 3 sample (Table 3.3).

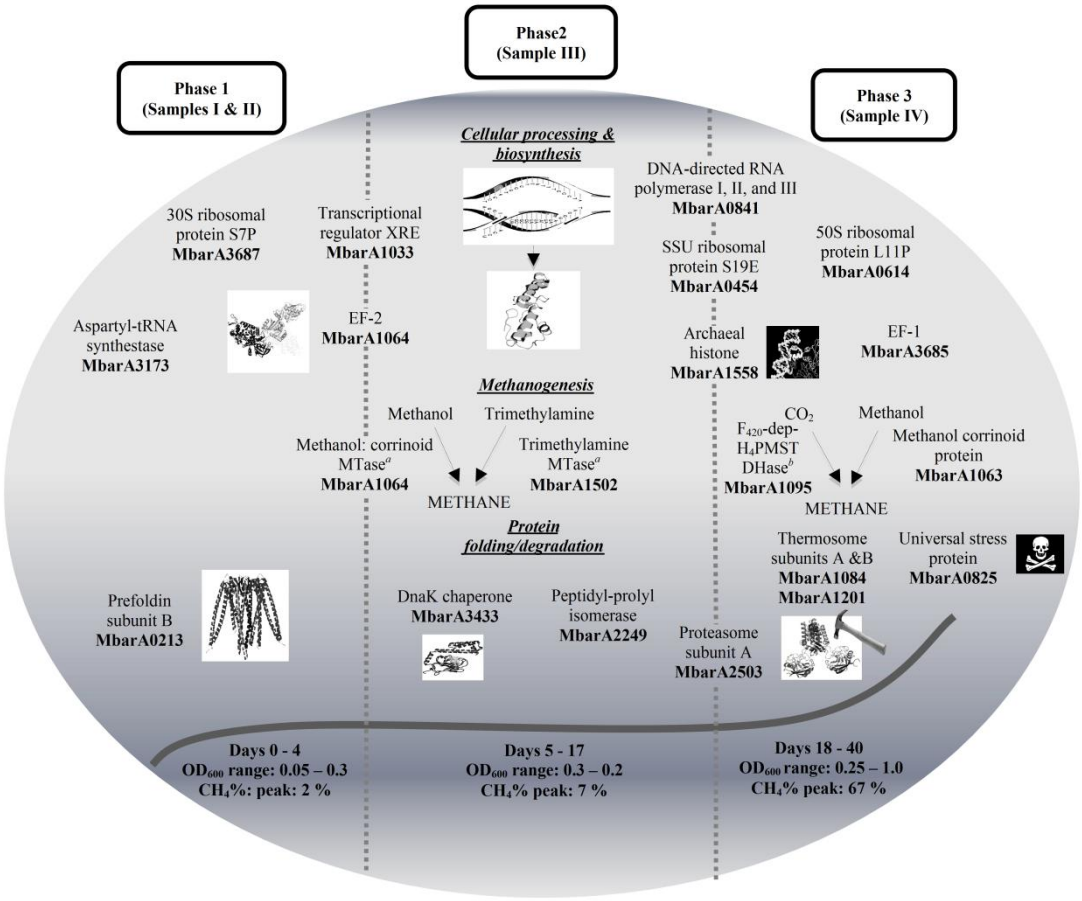


Fig. 3.3 Differentially expressed proteins at specific phases of *M*. *barkeri* growth at 15°C. Proteins located on broken line indicative of conserved result between particular phases. [a]MTase: methyltransferase, [b]DHase: dehydrogenase.

Four proteins relating to transcription and translation were upregulated in Phase 2 and Phase 3 samples when compared with Phase 1 samples (Table 3.3). These included a protein involved in transcription (MbarA0841) and a protein biosynthesis catalyst (MbarA3685). However, EF-2 protein (MbarA1064) as well as transcriptional regulator (MbarA1033) were found to be upregulated in Phase 1 and Phase 2 samples when compared with Phase 3 sample (Table 3.3). This result was consistent with the LTSR study where both of these proteins were detected at 15°C, suggesting that they might be important for initial stages of low-temperature adaptation. Two other DNA binding and repair proteins were found to be upregulated in Phase 2 and Phase 3 samples when compared to Phase 1 samples. These included the DNA regulator MbarA1558, involved in transcription (Weidenbach *et al*., 2008). General metabolism proteins identified in the LTAS study were found to be involved in diverse pathways, including amino acid and vitamin biosynthesis, glycolysis and electron transport (Table 3.3). Six of these proteins, which included the oxidoreductase protein MbarA2189, essential in glycolysis and biosynthesis of secondary metabolites, were expressed at a higher level in Phase 3 sample when compared to Phase 1 samples.

Proteins involved in oxidative stress, protein folding and proteolysis were also represented in the LTAS data set (Table 3.3). Two thermosome proteins (MbarA1084 and MbarA1201) were upregulated in the Phase 3 sample when compared with Phase 1 samples, which was also evident for a universal stress protein (MbarA085; Fig. 3.3). Also, a peptidyl prolyl isomerase (Mbar A2249) was upregulated in the Phase 2 sample when compared to other samples. Although little is known regarding chaperone-like activity of peptidyl prolyl isomerase, a study on *Methanococcoides burtonii* reported the upregulation of this protein when grown at 4°C compared to 26°C (Goodchild *et al*., 2005), therefore suggesting this protein may play an important role in low-temperature adaptation of *M. barkeri*. Finally, a DnaK heat shock protein (MbarA3433) was found to be upregulated in Phase 2 and a proteolysis protein (MbarA2503) was found to be upregulated in the Phase 3 sample when compared to Phase 1 samples (Fig. 3.3, Table 3.3).

In order to complement the iTRAQ results, 2-DGE was undertaken on *M.barkeri* samples grown at 37°C and 15°C (LTAS study). These samples were taken at time-points with similar OD_{600} and CH_4% levels for each temperature (Table 3.1). An average of 179 (SD 18.4; n=6) reproducible protein spots were detected on 2-D gels, from which 79 (44%) were found to be upregulated at 37°C, while 39 (22%) proteins were upregulated at 15°C and 61 (34%) were conserved for each temperature. Two proteins representing differential expression for each temperature were excised and identified using nLC-ESI-MS/MS. The universal stress protein (MbarA0825) was found to be upregulated (up 4.5-fold) at 37°C, while a proteasome (MbarA2503) was expressed at a higher level in the 15°C sample (up 2.9-fold) (Fig. 3.4). In the LTSR iTRAQ experiment, the universal stress protein was upregulated at 37°C while the proteasome was upregulated at 15°C, thus showing consistency between the two proteomic techniques applied (Table 3.3). A protein responsible for methyl coenzyme M reduction (MbarA0893) was conserved for each temperature, which was in agreement with the LTSR iTRAQ experiment also.

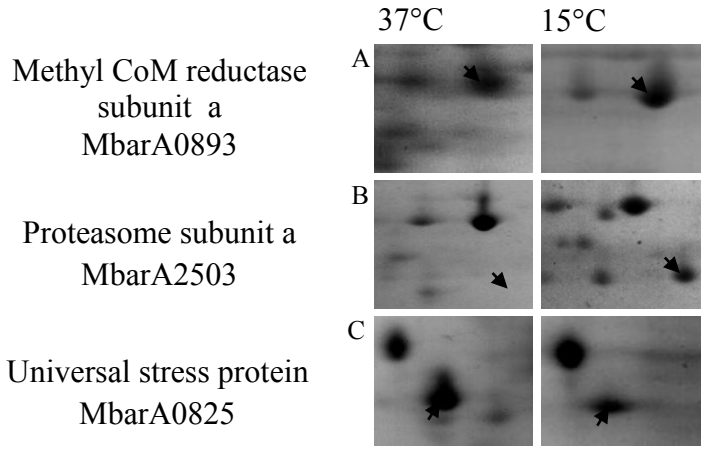


Fig. 3.4. **2-DGE gel sections comprising proteins extracted from *M. barkeri* grown at 37°C and 15°C.** (A) Methyl coenzyme M reductase subunit G protein (MbarA0893) conserved for each temperature (37°C/15°C, up 1.24-fold) (B) Proteasome subunit A (MbarA2503) upregulated at 15°C (15°C/37°C, up 2.87-fold) (C) Universal stress protein (MbarA0825) upregulated at 37°C (37°C/15°C, up 3.81-fold).

3.4 Discussion

The low-temperature adaptation of *M. barkeri* was investigated through two experimental approaches: (A) low-temperature shock response (LTSR) and (B) low-temperature adaptation strategies (LTAS). In the LTSR experiment, we applied an iTRAQ method to characterise the protein expression profile of *M. barkeri* before and after a temperature drop from 37°C to 15°C. In the LTAS experiment, *M. barkeri* growth at 15°C was characterised through proteomic samples taken at specific time-points to coincide with a particular phase of growth. Also, cell viability measurements were recorded, which included measurements at the proteomics sampling time-points, in order to give a concise overview of the sub-mesophilic growth capacity of *M. barkeri*. For the LTSR experiment, a total of 20 (13%) proteins were found to be upregulated at either 37°C or 15°C, while in the LTAS experiment there were 43 (41%) proteins found to be differentially expressed during incubation at 15°C (Table 3.3). Overall, this study provided new insights into the low-temperature adaptation of *M. barkeri*.

Methanogenesis from methanol was found to be significant for *M. barkeri* at 15°C, with corrinoid protein (MbarA3637) upregulated 2.8-fold at this low-temperature when compared with 37°C. During growth at 15°C, two methanol metabolism proteins (MbarA1063 and MbarA1064) were found to be upregulated at different time-points. Taken together, these results indicate the preference of *M. barkeri* to utilise methanol in sub-mesophilic conditions. Two H_2/CO_2 metabolism proteins were found to be upregulated at 37°C (MbarA1095, up 3.3-fold, MbarA1763, up 3.2-fold), which might suggest a preference for H_2/CO_2 at mesophilic temperatures. The pathway involved in trimethylamine and dimethylamine metabolism was found to be active in *M. barkeri* during growth at 15°C. Interestingly, a methyltransferase protein (MbarA1502), which catalyses trimethylamine reduction to methyl-coenzyme M, was upregulated during the stationary phase when compared to other stages of growth (Table 3.3). Thus, the ability to undertake methylamine methanogenesis at 15°C may confer an advantage for *M. barkeri* survival in a low-temperature environment.

Proteins involved in translation were detected in both iTRAQ experiments. The EF-2 protein MbarA1064 was upregulated at 15°C when compared to 37°C, and furthermore, it was found to be upregulated in the initial phases of growth at 15°C (Phase 1 and Phase 2 samples when compared to Phase 3 sample; Fig. 3.3). Therefore, this protein may play a role in the initial stages of low-temperature adaptation in *M. barkeri*. EF-2 has been investigated in relation to the psychrophylic *Methanococcoides burtonii*, where it was found to be highly flexible (unstable) in low-temperature growth, thus resulting in high enzymatic activity (Thomas & Cavicchioli, 1998), thus it is conceivable that EF-2 flexibility may confer an advantage for sub-mesophilic growth of *M. barkeri*. Also, the importance of low-temperature DNA binding and repair was evident through the upregulation of XRE transcriptional regulator (MbarA1033) as well as two other DNA binding proteins found at 15°C.

Cellular chaperone proteins were predominant at 37°C. A chaperonin protein (MbarA1543) was upregulated at 37°C, which correlated with the previous finding of up-regulation of the corresponding RNA transcript under heat stress conditions (Zhang *et al*., 2006). In addition, a thermosome (Ths) protein (MbarA1084) was found to be upregulated at 37°C. This Ths protein was also found to be upregulated in the latter phase of growth at 15°C (Phase 3 sample; Fig. 3.3), suggesting an important functional role during sub-mesophilic exponential growth conditions. A DnaK chaperone (MbarA3433), which was found to be upregulated in the stationary phase (Phase 2), has been recorded to play an important role in cell survival during cell stress conditions (Clarens *et al*., 1995), particularly heat stress (De Biase *et al*., 2002). The recombinant expression of DnaK from the psychrophilic *Shewanella sp*. in *E.coli* was found to confer the ability to grow at 15°C to an *E.coli* mutant strain (Yoshimune *et al*., 2005). Together with the peptidyl propyl isomerase (MbarA229), these proteins are important for low-temperature adaptation in *M.barkeri* and require further examination to characterise their respective roles in more detail.

Cell viability measurements recorded during growth of *M. barkeri* at 15°C gave insights into the capacity of this methanogen to survive in sub-mesophilic

conditions. This was highlighted with *M. barkeri* cells showing 83% cell viability after 17 days of incubation at 15°C. Together with proteomic data, these findings may contribute to a better understanding of *Methanosarcina* survival in low-temperature environments such as LTAD. It is conceivable that low-temperature substrate utilisation other than acetate takes place in such systems and thus may offer a favourable advantage by bypassing direct competition for acetate. Targeted functional studies are required to uncover the metabolic and adaptation strategies of *Methanosarcina spp*. in a mixed community context. The application of a polyphasic experimental design would be ideal, combining for example DNA quantification and proteomics methods with isotope labelling studies (e.g. separate incubations with C^{14} labelled methanol, H_2/CO_2 and acetate). Together these approaches may shine some light on the functional role of his methanogen in LTAD systems.

3.5 Conclusions

This mesophilically characterised methanogen was capable of growth and methane production in sub-mesophilic conditions. Proteomic investigation gave an indication that a preference for methanol metabolism may be occurring during low-temperature growth, with the upregulation of many key metabolic proteins during growth at 15°C in comparison to 37°C, for example, in the LTSR study, two H_2/CO_2 metabolism proteins were upregulated at 37°C, with the methanol corrinoid protein MbarA3637 upregulated at 15°C. Cell viability measurements were a key addition to growth experiments, as it was observed that at low-temperature growth, low cell numbers correlated with high cell viability. Also, it may be possible that the translational protein EF-2 may confer a psychrotolerant advantage to *M. barkeri* during low-temperature growth, with this protein being upregulated at 15°C compared with 37°C growth in the LTSR study, while also being upregulated during early phase growth at 15°C during the LTAS experiment. Further research is required to functionally characterise this psychrotolerant methanogen in a mixed community context, which will give further insights into the role this organism has in low-temperature environments such as in LTAD biofilms.

3.6 References

Abram, F., Enright, A.M., O'Reilly, J., Botting, C.H., Collins, G. and O'Flaherty, V. (2011) A metaproteomic approach gives functional insights into anaerobic digestion. *J. Appl. Microbiol.* **110**, 1550-1560.

APHA. (1998) Standard methods for the examination of water and wastewater In: Clesceri, L.S., Greenberg, A.E., Eaton, A.D (Eds.), Washington, DC, U.S.A.

Battin, T.J., Willie, A., Sattler, B. and Psenner, R. (2001) Phylogenetic and functional heterogeneity of sediment biofilms along environmental gradients in a glacial stream. *App. Environ. Microbiol.* **67**, 7099–7807.

Bock, A.K. and Schonheit, P. (1995) Growth of *Methanosarcina barkeri* (Fusaro) under non-methanogenic conditions by the fermentation of pyruvate to acetate: ATP synthesis via the mechanism of substrate level phosphorylation. *J. Bacteriol.* **177**, 2002-2007.

Buchenau, B. and Thauer, R.K. (2004) Tetrahydrofolate-specific enzymes in Methanosarcina barkeri and growth dependence of this methanogenic archaeon on folic acid or p-aminobenzoic acid. *Arch. Microbiol.* **182**, 313-325.

Cavicchioli, R. (2006) Cold adapted archaea. *Nat. Rev. Microbiol.* **4**, 331–343

Church, M.J., DeLong, E.F., Ducklow, H.W., Karner, M.B., Preston, C.M. and Karl, D.M. (2003) Abundance and distribution of planktonic archaea and bacteria in the waters west of the Antarctic Peninsula. *Limnol.Oceanogr.* **48**, 1893-1902.

Clarens, M., Macario, A.J.L. and Conway de Marcario, E. (1995) The archaeal dnaK-dnaJ gene cluster: organisation and expression in the methanogen *Methanosarcina mazei*. *J. Mol. Biol.* **250**, 191-201.

Collins, G., Enright, A.M., Scully, C., Mahony, T. and O'Flaherty, V. (2006) Application and biomolecular monitoring of psychrophilic anaerobic digestion. *Water Sci. Technol.* **54**, 41-47.

Connaughton, S., Collins, G. and O'Flaherty, V. (2006) Psychrophilic and mesophilic anaerobic digestion of brewery effluent: a comparative study. *Water Res*. **40**, 2503-2510.

De Biase, A., Marcario, A.J. and Conway de Marcario, E. (2002) Effect of heat stress on promoter binding by transcription factors in the cytosol of the archaeon *Methanosarcina mazeii*. *Gene*. **282**, 189-197.

Diaz, E.E., Stams, A.J., Amils, R. and Sanz, J.L. (2006) Phenotypic properties and microbial diversity of methanogenic granules from a full-scale upflow anaerobic sludge bed reactor treating brewery wastewater. *Appl*. *Environ*. *Microbiol*. **72**, 4942-4949.

Feist, A.M., Scholten, J.C.M., Palsson, B.O., Brockman, F.J., and Ideker, T. 2006. Modelling methanogenesis with a genome-scale metabolic reconstruction of Methanosarcina barkeri. Mol. Syst. Biol. 2:2006.0004.

Fey, A. and Conrad, R. (2000) Effect of temperature on carbon and electron flow and on the archaeal community in methanogenic rice field soil. *Appl*. *Environ*. *Microbiol*. **66**, 4790-479.

Galand, P.E., Lovejoy, C. and Vincent, W.F. (2006) Remarkably diverse and contrasting archaeal communities in a large arctic river and the coastal Arctic Ocean. *Aqua*. *Microb*. *Ecol*. **44**, 115-126.

Goenrich, M., Duin, E.C., Mahlert, F. and Thauer, R.K. (2005) Temperature dependence of methyl-coenzyme M reductase activity and of the formation of the methyl-coenzyme M reductase red2 state induced by coenzyme B. *J*. *Biol*. *Inorg*. *Chem*. **10**, 333–342.

Goodchild, A., Raftery, M., Saunders, N.F., Guilhaus, M. and Cavicchioli, R. (2005) Cold adaptation of the Antarctic archaeon, *Methanococcoides burtonii* assessed by proteomics using ICAT. *J*. *Proteome Res*. **4**, 473-480.

Griffin, M.E., McMahon, K.D., Mackie, R.I. and Raskin, L. (1998) Methanogenic population dynamics during start-up of anaerobic digesters treating municipal solid waste and biosolids. *Biotechnol*. *Bioeng*. **57**, 342-355.

Hao, X., Wang, Q., Zhang, X., Cao, Y. and van Loosdrecht, M.C.M. (2009) Experimental evaluation of decrease in bacterial activity due to cell death and activity decay in activated sludge. *Water Res*. **43**, 3604–3612.

Hoj, L., Olsen, R.A. and Torsvik, V.L. (2008) Effects of temperature on the diversity and community structure of known methanogenic groups and other archaea in high Arctic peat. *ISME J*. **2**, 37-48.

Liu, Y., Xu, H., Show, K. and Tay, J. (2002) Anaerobic granulation technology of wastewater treatment. *Microbiol*. *Biotechnol*. **18**, 99-113.

Maeder, D.L., Anderson, I., Brettin, T.S., Bruce, D.C., Gilna, P., Han, C.S., Lapidus, A., Metcalf, W.M., Saunders, E., Tapia, R. and Sowers, K.R. (2006) The *Methanosarcina barkeri* genome: comparative analysis with *Methanosarcina acetivorans* and *Methanosarcina mazei* reveals extensive rearrangement within methanosarcinal genomes. *J*. *Bacteriol*. **188**, 7922–7931.

Mazumder, T.K., Nishio, N., Fukuzaki, S. and Nagai, S. (1986) Effect of Sulfur-Containing Compounds on Growth of *Methanosarcina barkeri* in Defined Medium. *Appl*. *Environ*. *Microbiol*. **52**, 617-622.

McHugh, S., Carton, M., Collins, G. and O'Flaherty, V. (2004) Reactor performance and microbial community dynamics during anaerobic biological treatment of wastewater at 16-37°C. *FEMS Microbiol*. *Ecol*. **48**, 369-378.

McHugh, S., Collins, G. and O'Flaherty, V. (2006) Long-term, high-rate anaerobic biological treatment of whey wastewaters at psychrophilic temperatures. *Biores*. *Technol*. **97**, 1669-1678.

McKeown, R., Scully, C., Enright, A.M., Chinalia, F.A., Lee, C., Mahony, T., Collins, G. and O'Flaherty, V. (2009) Psychrophilic methanogenic community development during long-term cultivation of anaerobic granular biofilms. *ISME*. *J*. **3**, 1231-1242.

Morozova, D. and Wagner, D. (2007) Highly resistant methanogenic archaea from Siberian permafrost as candidates for the possible life on Mars. *International J*. *Astrobiol*. **6**, 68-69.

Muller, V., Blaut, M. and Gottschalk, G. (1986) Utilization of methanol plus hydrogen by *Methanosarcina barkeri* for methanogenesis and growth. *Appl*. *Environ*. *Microbiol*. **52**, 269-274.

Nomura, T., Nagao, T., Yoshihara, A., Tokumoto, H. and Konishi, Y. (2008) Selective immobilisation of acetoclastic methanogens to support material. *J*. *Soc*. *Powd*. *Technol*. **43**, 653-659.

O'Reilly, J., Lee, C., Collins, G., Chinalia, F.A., Mahony. T. and O'Flaherty, V. (2009) Quantitative and qualitative analysis of methanogenic communities in mesophilically and psychrophilically cultivated anaerobic granular biofilms. *Water Res*. **43**, 3365-3374.

Pernthaler, J., Glockner, F.O., Unterholzner, S., Alfreider, A., Psenner, R. and Amann, R. (1998) Seasonal community and population dynamics of pelagic bacteria and archaea in a high mountain lake. *Appl*. *Environ*. *Microb*. **64**, 4299–4306.

Shilova, I.V., Seymour, S.L., Patel, A.A., Loboda, A., Tang, W.H., Keating, S.P., Hunter, C.L., Nuwaysir, L.M. and Schaeffer, D.A. (2007) The paragon algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Molec*. *Cell*. *Prot*. **6**, 1638-1655.

Siggins, A., Enright, A.M. and O'Flaherty, V. (2011) Methanogenic community development in anaerobic granular bioreactors treating trichloroethylene (TCE)-contaminated wastewater at 37°C and 15°C. *Water Res*. **45**, 2452-2462.

Thomas, T. and Cavicchioli, R. (1998) Archaeal cold-adapted proteins: structural and evolutionary analysis of the elongation factor 2 proteins from psychrophilic, mesophilic and thermophilic methanogens. *FEBS Lett*. **439**, 281-286.

Veit, K., Ehlers, C., Ehrereich, A., Salmon, K., Hovey, R., Gunsalus, R.P., Deppenmeier, U. and Schmitz, R.A. (2006) Global transcriptional analysis of *Methanosarcina mazei* strain GÖ1 under different nitrogen availabilities. *Molec*. *Gen*. *Genom*. **276**, 41-55.

Wagner, D., Kobabe, S., Pfeiffer, E.M. and Hubberten, H.W. (2003) Microbial controls on methane fluxes from a polygonal tundra of the Lena Delta, Siberia. *Perm. Periglac. Process.* **14**, 173-185.

Weidenbach, K., Ehlers, C., Kock, J., Ehrenreich, A. and Schmitz, R.A. (2008) Insights into the NrpR regulon in *Methanosarcina mazei* Gö1. *Arch. Microbiol.* **190**, 319-332.

Yoshimune, K., Galkin, A., Kulalova, L., Yoshimura, T. and Esaki, N. (2005) Cold-active DnaK of an Antarctic psychrotroph *Shewanella* sp. Ac10 supporting the growth of dnaK-null mutant of *Escherichia coli* at cold temperatures. *Extremoph.* **9**, 245-250.

Zhang, W., Culley, D.E. and Brockman, F.L. (2006) DNA microarray analysis of anaerobic *Methanosarcina barkeri* reveals responses to heat shock and air exposure. *J. Ind. Microbiol. Biotechnol.* 33, 784-790.

Zhu, J., Zheng, H., Guomin, A., Zhang, G., Liu, D., Xiaoli, L. and Dong, X. (2012) The genome characteristics and predicted function of methyl-group oxidation pathway in the obligate acetoclastic methanogens, *Methanosaeta* spp. *PLoS ONE.* **7**, e36756.

Chapter 4

Characterisation of microbial communities during low-temperature anaerobic digestion using a polyphasic functional approach

A condensed version of this chapter is being prepared for submission to the ISME Journal:

Gunnigle, E., Nielsen, J.L., Fuszard, M., Botting, C.H., Abram, F., O'Flaherty, V. (2012). Characterisation of microbial communities during low-temperature anaerobic digestion using a polyphasic functional approach. Submission: Jan 2013.

Abstract

As low-temperature anaerobic digestion (LTAD) is a complex microbial mediated process, a clear understanding of the community-based interactions and succession of microbial groups is required if the true potential of this technology is to be realised for full-scale application. This study employed PCR-based (clone libraries, qPCR, DGGE [based on both RNA and DNA]) and PCR-independent (specific methanogenic activity [SMA] profiling, microautoradiography fluorescent *in situ* hybridisation [MAR-FISH], and metaproteomics) approaches to link community structure with the key functional groups underpinning the granular biomass sampled. Duplicate AD bioreactors (R1 & R2) were operated at 37°C for 125 days, with a temperature drop to 15°C, and operation for an additional 115 days, with biomass samples taken throughout the trial and studied using the various methodologies applied. Each bioreactor treated a synthetic volatile acid-based wastewater with an organic loading rate (OLR) of 3 kg COD m^{-3} d^{-1}. Clone library analysis indicated a predominance of the *Proteobacteria* group in biomass at both temperatures (48% [R1] and 62% [R2] of bacterial clones on day 125[37°C], with 45% [R1] and 62% [R2] on day 240[15°C]). This bacterial group was also represented in metaproteomic data, with 14 (82%) of the bacterial proteins identified assigned to this group. These included an acetyl CoA synthetase protein (Syn_02635), upregulated in low-temperature biomass. The *Methanosaeta* were prominent amongst the archaea at both temperatures, as confirmed through PCR-based (> 98% of total measured methanogenic population in qPCR analysis) and PCR-independent (27 [64%] of archaeal proteins identified) approaches. A F_{420} hydrogenase protein (Mcon_2320) was assigned to this methanogenic group, with a possible role in energy conservation at low-temperatures. MAR-FISH confirmed an increase in hydrogenotrophic activity in conjunction with temperature decrease, through an increase in specific radiolabelled substrate uptake. Other key findings included the identification of methanogenic functional groups through RNA-derived DGGE analysis such as *Methanospirillum*, which was not detected through DNA-based clone library analysis.

4.1 Introduction

Water shortage is a global issue. As a result, many people drink and use water far below the standard required. It is estimated that over 3 million people die in developing and developed countries each year from waterborne diseases like diarrhea, which could be easily prevented by the efficient treating the water (WHO, 2008). As such, many countries have set out targets within a water framework directive to develop and apply efficient water treatment systems. The development of high-rate anaerobic digestion (AD) wastewater treatment has provided a low-cost low-energy alternative to traditional aerobic systems, offering a sustainable option moving forward in improved water management and environmental protection. Recent attention has focused on applying this technology in a low-temperature context. This has been facilitated by advances in reactor design (e.g. the combination of granular and anaerobic filter [AF] configurations; the incorporation of membrane separations) and an improved understanding of the physical, chemical and biological challenges presented, thus enabling AD to be applied at sub-mesophilic temperatures (Verstraete & Vandedivere, 1999; Dohanyis *et al*., 2004; Martinez-Sosa *et al*., 2011).

As AD is a microbially-mediated process, the succinct understanding of the organisms underpinning these systems is essential for the improved optimisation of low-temperature anaerobic digestion (LTAD), for example, to facilitate deterministic modelling of the system (Ramirez *et al*., 2009). Indeed, LTAD has been developed using an expanded granular sludge bed (EGSB) bioreactor design, which allows high mixing intensities (mass transference; Kato *et al*., 1997) and the retention of psychroactive biomass (McKeown *et al*., 2009). The microbial consortia functioning in LTAD have been investigated at laboratory scale, with key insights gathered on community structure (Enright *et al*., 2009; Syutsubo *et al*., 2008). In these studies, a structural shift was recorded, with the selection of psychroactive organisms evident during the acclimation process to sub-mesophilic temperature operation.

Several LTAD studies have documented cold-mediated succession in microbial communities to be underpinned by efficient methanogenic activity. For instance, a

study by Siggins *et al*. (2011), demonstrated the feasibility of treating a trichloroethylene-based wastewater at 7°C, with COD removal efficiency and biogas CH_4 content ≥74%, with hydrogenotrophic methanogens prevalent (71-89% of total methanogenic population throughout the trial). The increase in relative abundance of hydrogen utilising methanogens during LTAD has also been recorded in other studies, which includes a long-term bioreactor trial (McKeown *et al*., 2009), that found *Methanocorpusculum*-like organisms to comprise >50% of the archaeal clone library after ~ 1200 days of sub-mesophilic operation. However, what was also found in this study, as in other research (Collins *et al*., 2005; Enright *et al*., 2007; O'Reilly *et al*., 2009), was that acetoclastic methanogens were also prominent in low-temperature biomass, primarily *Methanosaeta*-like organisms, with this microbial group deemed important for the formation and structure of well-functioning granular sludge (McKeown *et al*., 2012). It has been proposed that homoacetogenesis is a key biochemical pathway during low-temperature methanogenesis in natural environments and, as such, increased acetoclastic methanogenic activity may arise from autotrophic actotogenesis (Kotsyurbenko, 2005). The reason for several conflicting reports of the routes of methanogenic carbon flow during LTAD of different wasteaters is still unresolved, which represents a key biological knowledge gap in relation to operation of LTAD systems.

Advancement in molecular research has provided opportunities to describe microbial dynamics in complex mixed microbial communities such as anaerobic granules. The application of metaproteomics, whereby the proteome from a specific sample is extracted and analysed, has been successfully applied in a LTAD context (Abram *et al*., 2011). Moreover, ecophysiological approaches, such as microautoradiography-fluorescent *in situ* hybridisation (MAR-FISH) have allowed discrete substrate uptake dynamics at a cellular level to to be realised, which in turn may uncover the important functional groups underpinning a dynamic system such as LTAD granular biofilms. These approaches linked with molecular identification methodologies, such as qPCR and DNA sequencing, hold great promise in unraveling the complex trophic interactions and, as such, improve our understanding in optimising AD application in low-temperature setups.

In this study, 16S rRNA gene-based techniques (clone libraries, qPCR, DGGE), in conjunction with metaproteomics and MAR-FISH, were employed to link community structure with the key functional groups underpinning granular biomass sampled during the operation of duplicate AD bioreactors initially operated at 37°C, with a direct temperature drop to 15°C. The primary focus of this study was methanogenic community succession and function in response to a change in environmental temperature.

4.2 Materials and Methods

4.2.1 Design and operation of EGSB bioreactors

A granular, anaerobic sludge was obtained from a mesophilic (37°C), full-scale (1500 m3) internal circulation (IC) alcohol wastewater treatment bioreactor at Carbery Milk Products, Ballineen, Cork, Ireland. Granules were of a regular spherical shape (Ø, c. 1-2 mm) with a volatile suspended solids (VSS) concentration of 96 g l^{-1} and were used to seed two laboratory (3.5l), expanded granular sludge bed (EGSB) bioreactors as described by Collins *et al*. (2004). Each bioreactor (R1 & R2) was used to treat of a synthetic volatile fatty acid-based wastewater consisting of acetic acid, propionic acid, butyric acid and ethanol with a chemical oxygen demand (COD) ratio of 1:1:1:1, to a total of 3g COD l^{-1}. The synthetic influent was buffered with $NaHCO_3$ and fortified, as described by Shelton & Tiedje, (1984) with macro- (10 ml l^{-1}) and micro- (1 ml l^{-1}) nutrients. During this study, both bioreactors were initially operated at 37°C for 125 days, with a direct temperature drop to 15°C for a further 115 days. The organic loading rate (OLR) applied to all bioreactors was 3 kg COD m^{-3} d^{-1} with a hydraulic retention time (HRT) of 24hr. Effluent was re-circulated through both systems at an applied up-flow velocity of 2.5 m h^{-1}.

4.2.2 Specific methanogenic activity testing

Seed biomass, and samples collected from the bioreactors on days 125 (37°C) and day 240 (15°C) were screened for metabolic capability using specific methanogenic activity (SMA) tests, which were performed using the pressure transducer technique (Colleran *et al*., 1992; Coates *et al*., 1996), in which, acetate (30 mM) and H_2/CO_2 (80:20, v/v) were employed as substrates in order to establish the activities of acetoclastic and hydrogenotrophic methanogens, respectively. Vials without any substrate, or with the addition of N_2/CO_2 (80:20, v/v) instead of H_2/CO_2 (80:20, v/v) in the case of hydrogenotrophic tests, served as controls.

4.2.3 Analytical methods

Reactor influent, effluent and biogas from all bioreactors were routinely sampled. Influent and effluent COD concentrations and % biogas CH_4 content were determined according to Standard Methods American Public Health Association (APHA, 1998), and % COD removal efficiency was calculated from influent and effluent measurements. Analysis of acetate VFA concentration of effluent samples taken during the trial were performed by heated (85°C) and agitated headspace, in a Varian Saturn 2000 GC/MS system, with CombiPAL auto-sampler (Varian Inc., Walnut Creek, CA). Separation was carried out on a Varian Capillary column, CP-WAX 58 (FFAP) CB (25 m length x 0.32 mm internal diameter x 0.2 μm film thickness, Varian). The injector volume was 2 ml and the injector temperature was maintained at 250°C. Helium was employed as the carrier gas, at a flow rate of 1 ml min^{-1}. The temperature program was as follows: 50°C (20 sec) to 110°C (20 sec) at a rate of 2°C min^{-1}; from 110°C to 200°C (20 sec) at a rate of 20°C min^{-1}. The MS-detector was operated in the scan mode in the range of 40-150 m z^{-1} at a temperature of 210°C. Identification of acetate was achieved by matching chromatographic retention times and spectra of standard compounds. Calibration curves of standard VFAs were constructed and used for relative concentration of acetate in effluent headspace samples and then expressed as mg l^{-1}.

4.2.4 Co-extraction of Genomic DNA and RNA

Genomic DNA and RNA from granular biomass sampled from R1 and R2 on Days 125, 126, 128, 132 and 240 was co-extracted by a phenol extraction method as described by Carrigg *et al*. (2007). Briefly, 500 mg of granular biomass from each bioreactor was crushed with a mortar and pestle prior to adding 500 μl of 1% cetyl trimethylammonium bromide (CTAB) extraction buffer (Griffiths *et al*., 2000) and 500 μl of phenol-chloroform-isoamyl alcohol (25:24:1; pH 8.0). The mixture was then lysed for 20 seconds with 500 mg of zirconia beads (0.5 and 0.1 mm), in a Mini Beadbeater-8 (Biospec) at the median speed setting. The samples were then centrifuged at 10,000 x *g* for 5 mins to separate the aqueous phase containing the nucleic acids (DNA & RNA) and then added into fresh RNase-free micro-centrifuge

tubes. Any trace of phenol was removed by adding equal volumes of chloroform-isoamyl alcohol (24:1), followed by centrifugation at 10,000 x *g* for 5 mins. Polyethylene glycol (PEG)-1.6M NaCl (30% w/v) was then added (2:1 v/v) to each sample to precipitate total nucleic acids at room temperature. Ice-cold ethanol (70% v/v) was used to wash samples prior to air-drying and resuspension with DEPC-treated water.

4.2.5 qPCR

Quantitative real-time PCR was performed on genomic DNA from days 125 (37°C), 128 (15°C) and 240 (15°C; Table 4.1) using a LightCycler 480 (Roche, Mannheim, Germany), with four methanogenic primer and probe sets, specific to two hydrogenotrophic orders (*Methanomicrobiales* and *Methanobacteriales)* and two acetoclastic families (*Methanosaetaceae* and *Methanosarcinaceae*), covering most methanogens present in anaerobic digesters (Lee *et al*., 2009; Yu *et al*., 2005; Appendix 1). All DNA samples were analysed with each primer and probe set in duplicate, as described in Chapter 2, Section 2.2.4. The volume-based concentrations (copies per µl template) were converted into the biomass-based concentration (copies per gram VSS) using the determined VSS concentration of each sludge sample used for the DNA extraction.

4.2.6 Clone library analysis of 16S rRNA gene

Bacterial and archaeal clone libraries from day 125 (37°C) and 240 (15°C) biomass were constructed (Table 4.1) with amplification of bacterial 16S rRNA genes with forward primer 27F (5' – AGA GTT TGA TCC TGG CTC AG – 3'; DeLong, 1992) and reverse primer 1392R (5' ACG GGC GGT GTG TRC – 3'; Lane *et al*., 1985). Amplification of archaeal 16S rRNA genes was achieved with forward primer 21F (5' – TTC CGG TTG ATC CYG CCG GA – 3'; Stackebrandt & Goodfellow, 1991) and reverse primer 958R (5' – YCC GGC GTT GAM TCC AAT T – 3'; DeLong, 1992). Primer details are outlined in Appendix 1. PCR master mix and reaction conditions were as described in Chapter 2, Section 2.2.3. Cloning (TOPO® XL),

phylogenetic ribosomal DNA restriction analysis (ARDRA), plasmid sequencing and phylogenetic analyses were performed as outlined in Chapter 2 section 2.2.3. The sequences generated by this study were deposited in the GenBank database under the accession numbers: KC145381-KC145410, K145412-KC145419, KC182519-KC182527.

4.2.7 Denaturing gradient gel electrophoresis (DGGE)

DGGE analysis of archaeal 16S rRNA genes was carried out on both DNA and RNA samples extracted from granular biomass taken on days 125, 126, 128, 132 and 240 (Table 4.1). Reverse transcriptase of total RNA was accomplished by a method adapted from Corgié *et al*. (2006). Briefly, the removal of DNA was achieved by incubating 5 µl of crude nucleic extracts at 37°C for 30 mins with 3U of RQ1 RNase-free DNase (Promega corp.) in buffer plus inhibitors of RNase (Recombinant Rnasin® Ribonucelase Inhibitor). Samples were then incubated with 1 µl of STOP DNase at 70°C for 15 mins to inactivate DNase and denature RNA for Reverse Transcription (RT). Control of DNA contamination was achieved by using an aliquot of each DNase treated sample in a standard 16S rRNA archaeal PCR, with a negative result confirming no DNA present. RT was then undertaken using 10 µl of DNase treated nucleic acid samples. The reaction mixture also comprised of 5X first strand buffer, 10 mM of each dNTP, 0.1 mM DTT, 1U/µl Recombinant Rnasin® Ribonuclease Inhibitor, 25 ng/µl random primers and 1U/µl of SuperScript™ III Reverse Transcriptase (Invitrogen). The RT reaction was accomplished using an iCycler (Biorad) with the following programme: 5 mins at 25°C, 50 mins at 50°C and 15 mins at 70°C. The generated cDNA was stored at -20°C prior to amplification. Initial PCR amplification of DNA and cDNA was undertaken using forward primer 787F (5' – ATTAG ATACC CSBGT AGTCC – 3') and 1059R (5' – GCCAT GCACC WCCTC T– 3'; Yu *et al*., 2005), with a 40-base pair GC clamp attached to the 5' terminus of the forward primer. The touchdown PCR program consisted of an initial denaturation at 94°C for 2 mins, followed by denaturation at 94°C for 30 seconds, annealing of primers (65°C-55°C; 1 cycle at 1°C increments; 20 cycles at 55°C) for 30 seconds and extension at 72°C for 30 seconds, followed by a final extension at 72°C for 10 mins. All PCR products were electrophoresed on 2%

agarose gels, and suitable amplicons were selected for DGGE analysis.

A 40µl aliquot of each GC-clamped PCR product was loaded onto a 10% (w/v) polyacrylamide gel containing a denaturing gradient of 40-60% (where 100% denaturant contained 7M urea, 40% formamide) and ran at 60°C and 70V for 16h in a D-Code system (BioRad, Hercules, CA). The DGGE gels were ethidium bromide stained for 10 mins and photographed under UV trans-illumination.

For sequencing and phylogenetic analysis, selected of bands were excised from DGGE gels using a sterile scalpel blade. The excised bands were suspended in 200 µl of sterile water, and stored at room temperature for three hours to elute DNA from the gel for use as a PCR template. These PCR reactions were performed under the conditions as described above and products were cloned using TOPO® TA (Invitrogen). Five resulting plasmids were extracted and a 2 µl aliquot was employed as a template for PCR using the same primers and conditions as described previously. For confirmatory purposes, PCR products were electrophoresed on a DGGE gel in parallel with the original samples. Bands obtained from extracted plasmids that underwent denaturation at the same gradient concentration as the original sample, and thereby migrated the same distance through the gel, were selected and sequenced (MWG).

Sequences from this study were aligned with 16S rRNA gene sequences retrieved from the BLASTn using Clustal X (Thompson *et al*., 1997), and GTR + gamma model DNA substitution implemented RAxML 7.0.3 programme was used for all phylogenetic analyses (Stamatakis, 2006). The resulting partial 16S rDNA gene sequences were deposited in the GenBank database under the accession numbers (KC305601- KC305623).

4.2.8 Microautoradiography- fluorescent in-situ hybridization (MAR-FISH)

MAR radiotracer incubations were undertaken as outlined by Andreasen & Nielsen, (1997). Aliquots of 2 ml (c. 2-4g$[VSS]^{-1}$) sludge granules from day 125 (37°C) and

day 240 (15°C) were incubated immediately after sampling. These incubations were undertaken anaerobically with either [^{14}C]-labelled acetate or sodium [^{14}C]-bicarbonate (American Radiolabelled Chemicals, Inc.). Twenty μCi (0.74 MBq) of radioactive substrate was added to each sample with [^{14}C]-labelled acetate samples supplemented with 2 mM unlabelled analogues to ensure radiolabelled substrate was not utilised straight away. In the same context, [^{14}C]-labelled sodium bicarbonate was pressurized with 0.5 bar (95 mV) H_2CO_2. Each sample (except for incubations with sodium [^{14}C]-bicarbonate) was pressurized with O_2-free N_2 gas. As a control for possible adsorption phenomena, autoclaved sludge granules were also incubated under the same conditions in parallel. Incubations were stopped at the predetermined time-points by fixing with paraformaldehyde (PBS; 130 mM sodium chloride and 10 mM sodium phosphate [pH 7.2]; final concentration 4% [w/v]) for 4 hrs at 4°C. Samples were washed ten times with 1X PBS to remove excess radioactive substrate and PFA. Liquid nitrogen was used to embed granules in OCT freezing medium prior to sectioning. Serial cryosections of 5-10 μm thickness were prepared, as described previously by Sekiguchi *et al*. (1999), and were immobilized on gelatin coated, acid –washed, coverslips (24 mm x 50 mm; VXR international).

Granular sections were then dehydrated and fluorescence *in situ* hybridization (FISH) was preformed as described by Schramm *et al*. (1998). The Cy3-labelled 16S rRNA probes used in this study related to the two functional methanogenic groups investigated. The acetoclastic group comprised of *Mx825* (*Methanosaeataceae*) and *SarcI551* (*Methanosarcinaceae*) probes, which were applied onto sections from [^{14}C]-labelled acetate incubations, while the hydrogenotrophic group comprised of *MB1174* (*Methanobacteriales*) and *MG1200b* (*Methanomicrobiales*) probes, which were applied onto sections from H_2, sodium [^{14}C]-bicarbonate incubations. The *Non388* probe (Wallner *et al*., 1993), complementary to the *Eub388* sequence, was used as a negative control.

The MAR procedure was undertaken following FISH. First, LM emulsion film (Kodak) was applied to all coverslips, which were then left to expose in complete darkness at 4°C for 4 days. After exposure, coverslips were developed by standard

photographic procedures (Andreasen & Nielsen, 1997). This involved placing the coverslips in the developer (Kodak D19; 40 g l^{-1}) for 2 min, following which, each coverslip preparation was washed in demineralised water for 1 min, and then fixed in 30% thiosulfate for 2 mins. Finally, the samples were washed for 2 mins in demineralised water and left to air dry. For microscopic analysis, air-dried coverslips were mounted in an anti-bleaching agent Citifluor (UKC Chemlab, UK). The coverslips were then placed upside down on glass slides, making it easier to visualize silver grains underneath the cells (Nielsen and Nielsen, 2005). An epifluorescent microscope (Carl Zeiss, Oberkochen, Germany) was used for detection of FISH signal, while light microscopy was used to access silver grain density and MAR-positive cells from the MAR procedure. Mar-positive cells were recorded as being a cell covered with more than five silver grains (Okabe *et al*., 2005). Fluorescence heat maps were also generated (ImageJ software) to compare corrected total cell fluorescence (CTCF; Sundar *et al*., 2012) of acetoclastic and hydrogenotrophic groups between the two sampling days (125 [37°C] and 240 15°C]).

2.2.9 Metaproteomics

Proteins were extracted from 50 ml granular sludge samples from R1 and R2 on days 125 (37°C), 128 (15°C) and 240 (15°C; Table 4.1), with 2-DGE protein separation undertaken using a protocol as described by Abram *et al*. (2011). Details of which can be found in Chapter 2, section 2.2.5.

Nanoflow liquid chromatography-electrospray ionization tandem mass spectrometry (nLC-ESI-MS/MS) and protein identification was carried out on proteins excised from the gels and subjected to in-gel digestion, using a ProGest Investigator in gel-digestion robot (Genomic Solutions) following standard protocols (Shevchenko *et al*., 1996) and as described in Chapter 2, section 2.2.6. MS/MS data for +1 to +5 charged precursor ions which exceeded 150 cps was processed using the Paragon™ search algorithm (Shilov *et al*, 2007) within ProteinPilot 4.0 software (ABSciex, Foster City, CA) against NCBInr database July 2012 (38496380 sequences) with no species restriction.

4.3 Results

4.3.1 Bioreactor performance

During the first 125 days of this study (PI-PIV; Table 4.1), high treatment efficiencies of the VFA-based wastewater were recorded for both R1 and R2. This was evident through a rapid start-up with mean % COD removal efficiencies of 91% (R1) and 90% (R2) achieved during initial 40 days of trial (PI), with biogas levels of 66% CH_4 (R1) and 63% CH_4 (R2; Table 4.1; Figure 4.1). In PII however, R2 showed a slight deviation from R1 with COD removal efficiency at days 50 and 68 dropping to 78% and 72%, respectively. The biogas quality of R2 was also low on these days, with 41.5% and 13.5% CH_4 recorded on those two days. This may be due to air bubbles, which were observed in recirculation tubing during, and between, these time-points. This perturbation may have disturbed specific microbial community members in R2 biomass, thus contributing to a reduction in treatment efficiency. Nevertheless, mean COD removal efficiencies recovered to >75% in both bioreactors for the remaining phases during operation at 37°C (Table 4.1; Figure 4.1).

After the applied temperature reduction to 15°C, no visible effect on bioreactor performance was initially observed (Table 4.1; Fig. 4.1). Conversely, R1 recorded an increase in COD removal efficiency from PV-PVI. However, there was a visible decline in R2 performance on days 128 and 132 with 37% and 32% % COD removal efficiency recorded. Interestingly, the biogas quality did not appear to be effected by the decrease in COD removal efficiency on these days, being maintained >60% throughout PV. The performance of R2 recovered well by PVI, with a mean COD removal efficiency of 68% being recorded. Both bioreactors exhibited stable treatment of VFA wastewater for the remainder of the trial, with R1 and R2 having COD removal efficiencies of >90% and biogas CH_4 levels of ≥60% at the end of the trial (Table 4.1; Fig. 4.1). Analysis of effluent acetate VFA concentrations identified a sharp increase of acetic acid in the effluent of R1 (0.49g COD l^{-1}) and R2 (0.47 g COD l^{-1}) on day 131, five days after the temperature was decreased to 15°C. (Fig. 4.2).

Table 4.1 Summary of the operational performance of the R1 and R2 reactors and the experimental approaches applied on samples taken at specific time-points during trial.* The seed biomass (day 0 sample) was also analysed by SMA profiling. Standard deviations are given in parenthesis

Temperature	**37°C**			**15°C**			
Operational periods	**PI**	**PII**	**PIII**	**PIV**	**PV**	**PVI**	**PVII**
Day(s)	0-40	41-80	81-125	126-132	133-160	161-200	201-240
Mean COD Removal (%)							
R1	93(3)	91(5)	91(6)	86(6)	88(6)	90(4)	92(2)
R2	91(3)	89(6)	90(5)	56(25)	68(18)	83(8)	88(3)
Mean CH_4 (%) in biogas							
R1	63(4)	60(4)	58(12)	56(8)	59(5)	58(6)	54(5)
R2	60 (13)	51(14)	57(6)	55(4)	55(10)	57(6)	55(6)

Sampling days	125	126	128	132	240
SMA*	●				●
Clone libraries	●				●
Q-PCR	●		●		●
DGGE	●	●	●	●	●
MAR-FISH	●				●
Metaproteomics	●		●		●

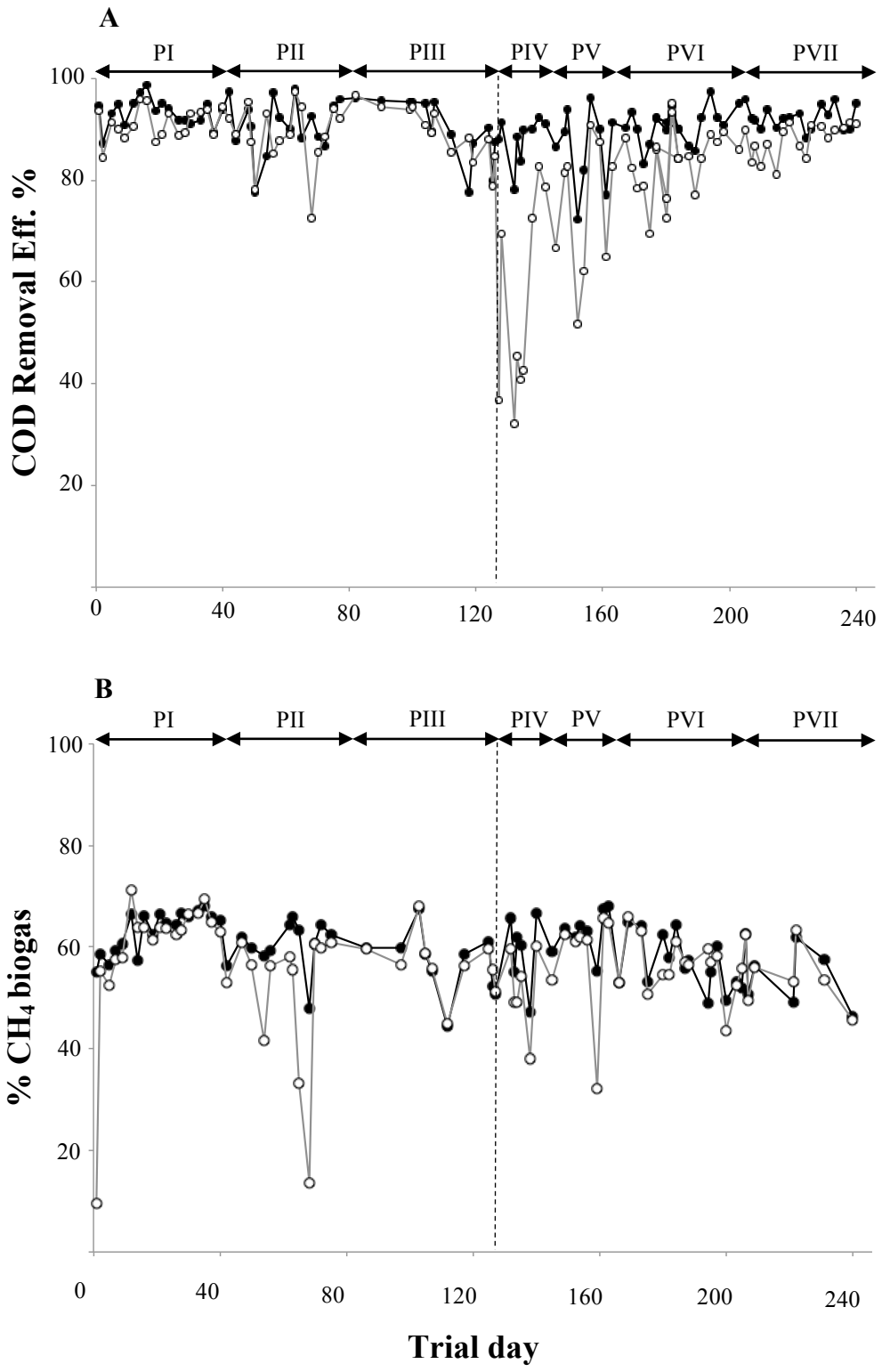


Fig. 4.1 Bioreactor performance including (**A**) % COD Removal efficiency (**B**) and % Biogas CH_4; R1 (●), R2 (○). Dashed line indicates point where operation temeprature was dropped to 15°C.

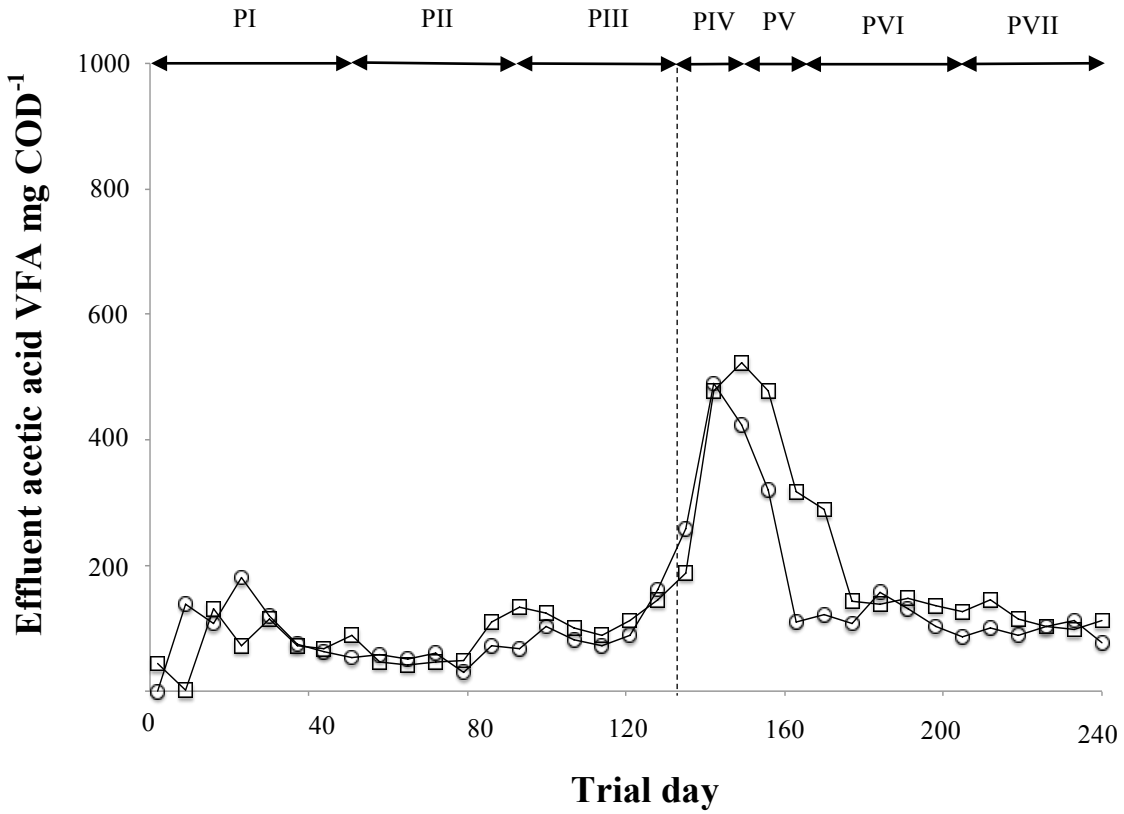


Fig. 4.2 Effluent acetic acid VFA for R1 (□) and R2 (○). Dashed line indicates point where operation temeprature was dropped to 15°C

4.3.2 Specific methanogenic activity of bioreactor biomass

The biomass used to seed R1 and R2 exhibited a higher specific methanogenic activity (SMA) with H_2/CO_2 as substrate, when compared to acetate, at both temperatures investigated (37°C and 15°C; Table 4.2). An increase was recorded in acetoclastic methanogenic activity in samples taken on day 125 for R1 and R2 of c.3-fold and c.2.6-fold, respectively, while SMA values at 15°C with acetate as substrates also increased compared to the seed biomass, with a c.2.5-fold increase noted in R1, for example. However, the values recorded for acetoclastic SMA, at 15˚C, on day 125 were much lower in comparison to 37°C results, with R2 only recording an 8% increase from the seed acetate SMA result (Table 4.2).

Meanwhile, an increase in SMA values with H_2/CO_2 as substrate was also evident on day 125, albeit not to the same extent as for acetoclastic methanogenesis. A c. 1.5-

hydrogenotrophic to acetatoclastic methanogenic activity.

Table 4.2 Specific methanogenic activity profiles of day 125 and day 240 biomass for direct methanogenic substrates, expressed as ml CH_4 $gVSS^{-1}d^{-1}$. All values are mean of triplicates ± std. (std. deviation /√n, n=3)

Biomass	Trial day	Temperature (°C)	Substrates	
			Acetate	H_2CO_2
Seed	0	37	176(16)	311(24)
	0	15	24(3)	40(5)
R1	125	37	516(73)	482(45)
	125	15	60(3)	43(15)
R2	125	37	453(48)	355(50)
	125	15	34(3)	46(1)
R1	240	37	892(36)	649(17)
	240	15	579(47)	489(62)
R2	240	37	814(14)	579(16)
	240	15	679(43)	514(19)

At the end of the trial, the biomass of both R1 and R2 exhibited a substantial increase in methanogenic activity against both substrates, at 15°C. Acetoclastic methanogenic activity still remained higher, with a c.10-fold and c.20-fold increase recorded in R1 and R2 on day 240 at 15°C. Interestingly, the acetate SMA was also higher at 37°C on day 240 when compared with day 125, with R1 having a c.1.7-fold increase, while R2 had a c.1.8-fold increase (Table 4.2).

Although to as great an extent as the acetatoclastic SMA, hydrogenotrophic methanogenic activity also exhibited an increase between day 125 and the end of the trial, with R1 recording a c.12-fold increase and R2 having an c.11-fold increase by day 240, at 15°C. As with acetoclastic SMA, the H_2-mediated activity at 37˚C on day was higher also when compared to day 125, with a c.1.3-fold and c.1.6-fold recorded

for biomass taken from R1 and R2, respectively (Table 4.2).

4.3.3 Bacterial 16S rRNA clone libraries

Thirty distinct operational taxonomic units (OTUs) were identified from the two bacterial clone libraries, from 384 clones analysed (Table 4.3). The relative abundance of clones present in each bioreactor biomass sample were assessed through amplified ribosomal rDNA restriction analysis (ARDRA) and sequencing. BLASTn search results and phylogenetic reconstruction revealed that bacterial clones related to the phylum *Proteobacteria* were dominant in both bioreactor biomass samples (Table 4.3; Fig. 4.3[A]; Fig. 4.4). Specifically, clones from day 125 biomass aligned closely with a *δ-proteobacteria* species, which represented 48% (R1) and 62% (R2) of day 125 clone library (Table 4.3; Fig. 4.4). This profile was also evident at the end of the trial, with eight OTUs closely aligned with *δ-proteobacteria* species, representing 45% of R1 bacterial clones and 62% of R2 clones.

Although clones relating to the phylum *Proteobacteria* were dominant in each bioreactor at both sampling points, changes in bacterial community structure was evident in relation to other bacterial groups. For example, in R2, there was an increase in *Synergistetes*-like clones in day 240 biomass with a single OTU comprising 14% of bacterial clone library in this sample. Also, there was an OTU closely aligned to a *Tenericutes* species, which was not detected in day 125 sample or in R1 biomass (Table 4.3; Fig. 4.4).

4.3.4 Archaeal 16S rRNA clone libraries

Representative clones from the 17 different OTUs identified in archaeal clone libraries grouped primarily with the order *Methanosarcinales*, with *Methanosaeta*-like clones comprising the majority of clones in both biomass samples (Table 4.4; Fig. 4.3[B]; Fig. 4.5). However in both R1 and R2 there was a decrease recorded in *Methanosaeta*-like clones, for example, in R2 day 125 biomass, clones closely

aligned to *Methanosaeta* comprised 64% of archaeal clone library, whereas in day 240 biomass, 45% of the library was assigned to this group (Fig. 4.5).

The order *Methanobacteriales* was represented in both bioreactor biomass samples, with levels remaining constant between sampling points, ranging from 10 to 13% in both bioreactors. Clones closely aligned with *Methanomicrobiales* were only evident in day 240 biomass samples, with *Methanolinea*-like clones representing 10% (R1) and 12% (R2) of archaeal clone library in these end point samples (Table 4.4; Fig. 4.5).

Finally, *Crenarchaeota*-like clones accounted for 11% (R1) and 21% (R2) of archaeal clone library in day 125 biomass. There was a slight increase of clones related to this phylum in day 240 samples. This comprised of a single OTU closely aligned with a *Staphylothermus* species, which represented 22% of the archaeal clone library in the R2 end-point sample on day 240 (Table 4.4; Fig. 4.5).

Table 4.3 Bacterial 16S rRNA gene sequence clones detected in clone libraries constructed from R1 and R2 bioreactor biomass on trial days 125 and 240. Value given is the number of clones per OTU, % cover of clone library is given in parenthesis. * indicates OTUs that are not relevant to a sample.

Library	Accession no.	R1 No. of clones	R2 No. of clones	Length (bp)	Closest relative (BLASTn)	Phylogenetic classification
Day 125	KC145381	88(46)	111(58)	1060	*Desulfovibrio alkalitolerans* (NR043069)	*Proteobacteria*
No. of clones analysed: 192	KC145382	15(8)	13(7)	1059	*Alkaliflexus imshenetskii* (NR042317)	*Bacteroidetes*
No. of unique OTUs: 11	KC145383	19(10)	4(2)	1130	*Levilinea saccharolytica* (NR040972)	*Chloroflexi*
	KC145384	6(3)	10(5)	1024	*Leptolinea tardivitalis* (NR040971)	*Chloroflexi*
	KC145385	4(2)	8(4)	746	*Desulfonatronum lacustre* (NR041848)	*Proteobacteria*
	KC145386	6(3)	*	779	*Aminobacterium colombiense* (NR027531)	*Synergistetes*
	KC145387	6(3)	8(4)	1084	*Levilinea saccharolytica* (NR040972)	*Chloroflexi*
	KC145388	27(14)	22(12)	1044	*Syntrophomonas zehnderi* (NR044008)	*Firmicutes*
	KC145389	*	2(1)	1082	Prolixibacter bellariivorans (NR043273	*Bacteroidetes*
	KC145390	11(6)	6(3)	1053	*Aminobacterium colombiense* (NR027531)	*Synergistetes*
	KC145391	10(5)	8(4)	970	*Levilinea saccharolytica* (NR040972)	*Chloroflexi*
Day 240	KC145392	31(16)	10(5)	1013	*Desulfuromusa ferrireducens* (NR043214)	*Proteobacteria*
No. of clones analysed: 196	KC145393	2(1)	4(2)	803	*Caldilinea aerophila* (N040878)	*Chloroflexi*
No. of unique OTUs: 19	KC145394	15(8)	23(12)	1023	*Syntrophobacter sulfatireducens* (NR043073)	*Proteobacteria*

KC145395	11(6)	27(14)	1040	*Aminobacterium colombiense* (NR027531)	*Synergistetes*
KC145396	2(1)	10(5)	1040	*Geobacter pickeringii* (NR043576)	*Proteobacteria*
KC145397	17(9)	2(1)	1042	*Syntrophomonas palmitatica* (NR041528)	*Firmicutes*
KC145398	2(1)	4(2)	977	*Levilinea saccharolytica* (NR040972)	*Chloroflexi*
KC145399	*	4(2)	949	*Anaeromyxobacter dehalogenans* (NR027547)	*Proteobacteria*
KC145400	6(3)	17(9)	818	*Aminiphilus circumscriptus* (NR043061)	*Synergistetes*
KC145401	29(15)	19(10)	1044	*Desulfovibrio gracilis* (NR044785)	*Proteobacteria*
KC145402	2(1)	21(11)	977	*Desulfuromonas alkaliphilus* (NR043709)	*Proteobacteria*
KC145403	25(13)	31(17)	1036	*Syntrophus gentianae* (NR029295)	*Proteobacteria*
KC145404	10(5)	*	994	*Levilinea saccharolytica* (NR040972)	*Chloroflexi*
KC145405	11(6)	6(3)	845	*Prolixibacter bellariivorans* (NR043273)	*Bacteroidetes*
KC145406	19(10)	8(4)	911	*Levilinea saccharolytica* (NR040972)	*Chloroflexi*
KC145407	4(2)	*	999	*Prolixibacter bellariivorans* (NR043273)	*Bacteroidetes*
KC145408	*	6(3)	978	*Acholeplasma morum* (NR042959)	*Tenericutes*
KC145409	4(2)	*	808	*Geobacter pickeringii* (NR043576)	*Proteobacteria*
KC145410	2(1)	*	1015	*Alkaliflexus imshenetskii* (NR042317)	*Bacteroidetes*

Table 4.4 Archaeal 16S rRNA gene sequence clones detected in clone libraries constructed from R1 and R2 bioreactor biomass on Trial Days 125 and 240. Value given is the number of clones per OTU, % cover of clone library is given in parenthesis. * indicates OTUs that are not relevant to a sample.

Library	Accession no.	R1 No. of clones	R2 No. of clones	Length (bp)	Closest relative (BLASTn)	Phylogenetic classification
Day 125	KC145412	19(10)	31(16)	943	*Ignisphaera aggregans* (NR043512)	*Crenarchaeota*
No. of clones analysed: 192	KC145413	140(73)	128(67)	689	*Methanosaeta concilii* (NR028242)	*Methanosarcinales*
	KC145414	4(2)	6(3)	915	*Methanobrevibacter smithii* (NR044786)	*Methanobacteriales*
No. of unique OTUs: 8	KC145415	6(3)	*	916	*Methanosaeta concilii* (NR028242)	*Methanosarcinales*
	KC145416	2(1)	10(5)	842	*Ignisphaera aggregans* (NR043512)	*Crenarchaeota*
	KC145417	6(3)	2(1)	914	*Methanosaeta concilii* (NR028242)	*Methanosarcinales*
	KC145418	12(6)	11(6)	924	*Methanobrevibacter smithii* (NR044786)	*Methanobacteriales*
	KC145419	4(2)	4(2)	915	*Methanobrevibacter smithii* (NR044786)	*Methanobacteriales*
Day 240	KC182519	4(2)	15(8)	904	*Methanobacterium formicicum* (NR025028)	*Methanobacteriales*
No. of clones analysed: 192	KC182520	6(3)	12(6)	926	*Methanosaeta harundinacea* (NR043203)	*Methanosarcinales*
	KC182521	4(2)	6(3)	915	*Methanosaeta concilii* (NR028242)	*Methanosarcinales*
No. of unique OTUs: 9	KC182522	10(5)	4(2)	790	*Methanomethylovorans hollandica* (NR028174)	*Methanosarcinales*
	KC182523	13(7)	23(4)	883	*Methanobacterium beijingense* (NR028202)	*Methanobacteriales*

KC182524	109(57)	81(42)	898	*Methanosaeta concilii* (NR028242)	*Methanosarcinales*
KC182525	4(2)	2(1)	900	*Methanobrevibacter smithii* (NR044786)	*Methanobacteriales*
KC182526	23(12)	41(22)	902	*Staphylothermus marinus* (NR044909)	*Crenarchaeota*
KC182527	19(10)	8(12)	909	*Methanolinea tarda* (NR028163)	*Methanomicrobiales*

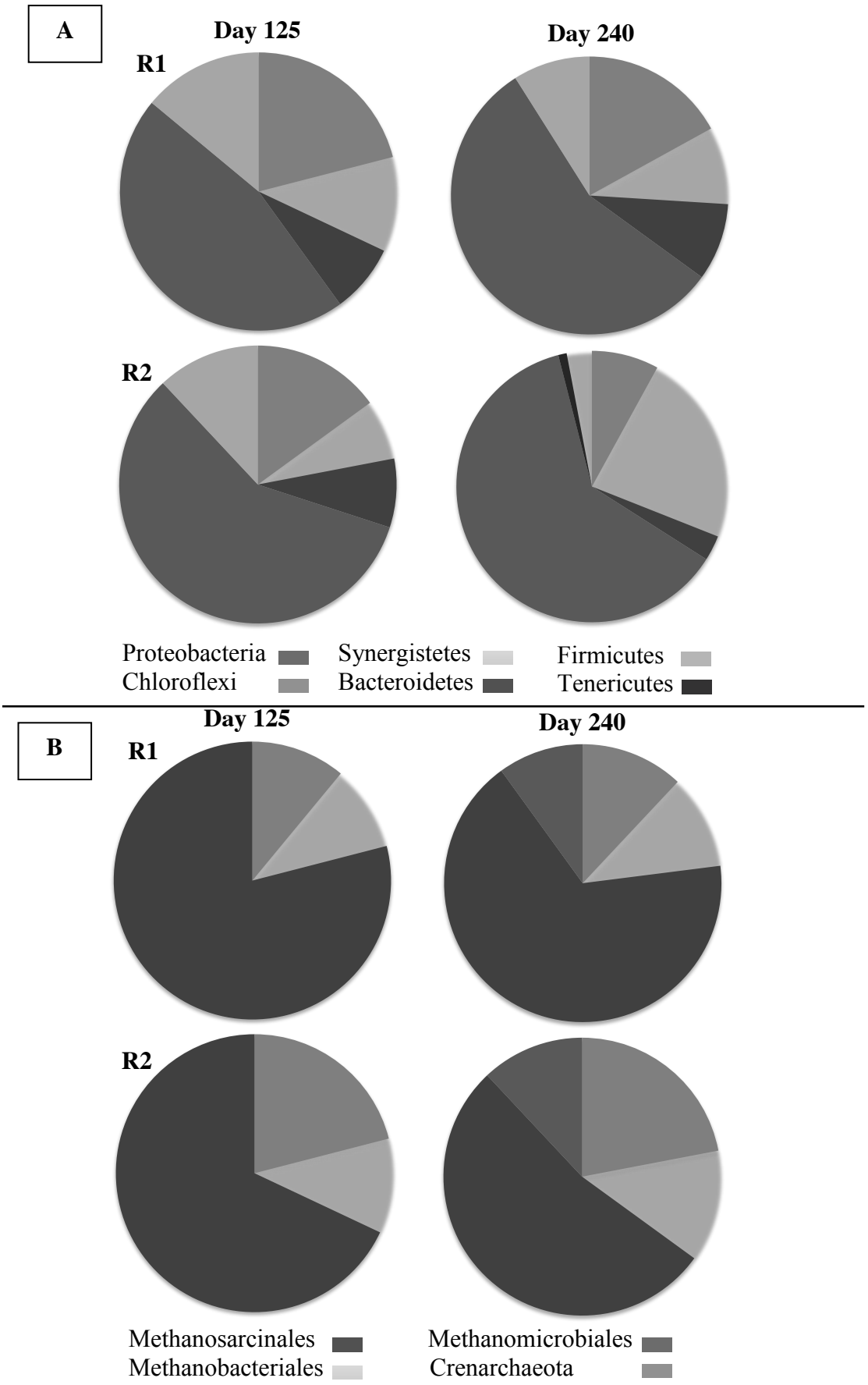


Figure 4.3 Comparative distribution of (**A**) bacterial and (**B**) archaeal groups detected by clone library analysis performed on bioractor biomass on days 125 and 240.

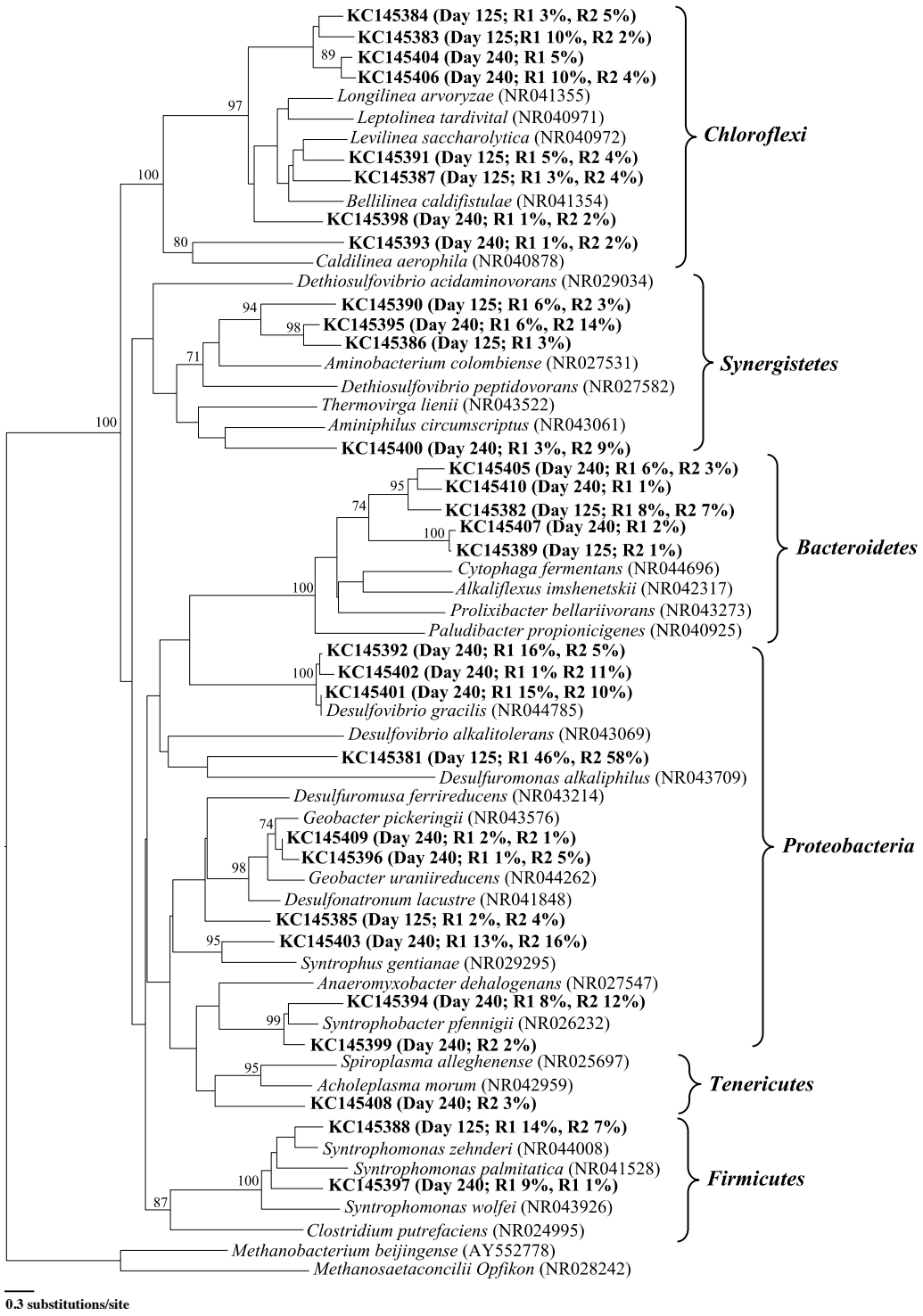


Figure 4.4 Phylogenetic affiliation of the bacterial 16S rRNA gene sequences identified from day 125 (37°C) and day 240 (15°C) bioreactor biomass, calculated using the GTR + gamma model DNA substitution implemented RAxML 7.0.3 (Stamatakis, 2006). Bootstrap replicates (total 1000 replicate samples) supporting the branching order are shown at relevant nodes.

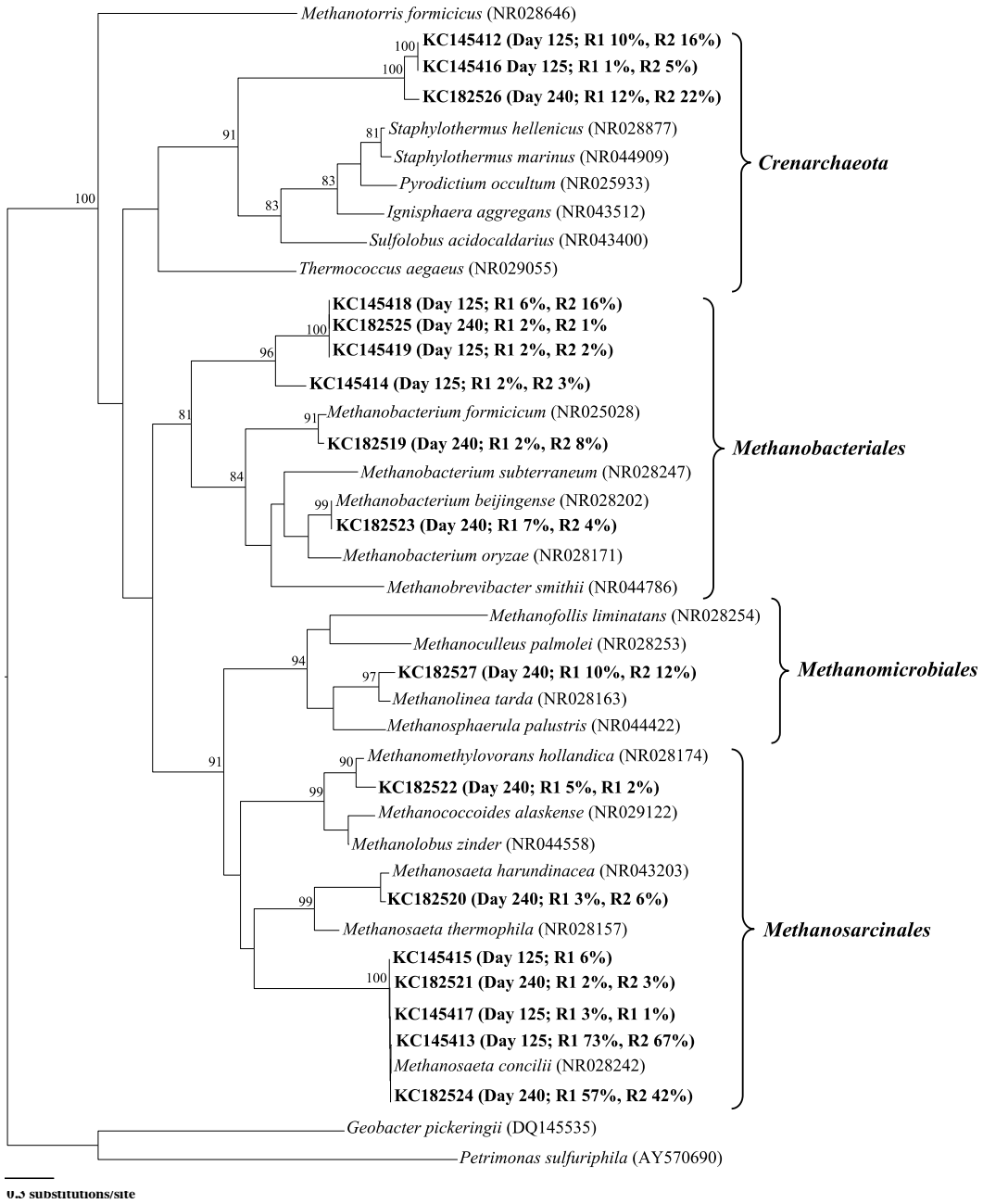


Figure 4.5 Phylogenetic affiliation of the archaeal 16S rRNA gene sequences identified from day 125 (37°C) and day 240 (15°C) bioreactor biomass, calculated using the GTR + gamma model DNA substitution implemented RAxML 7.0.3 (Stamatakis, 2006). Bootstrap replicates (total 1000 replicate samples) supporting the branching order are shown at relevant nodes.

4.3.5 qPCR

16S rRNA genes from the acetoclastic *Methanosaeteceae* were predominant in both bioreactors at each time-point (Fig. 4.6). In R1, there was a increase from 1.34 x 10^{11} to 4.33 x 10^{11} copies g VSS^{-1} between day 125 and the end-point sample on day 240, representing a c.2.6-fold increase. R2 relayed a similar result with a c.2.5-fold increase in *Methanosaetaceae* gene copie numbers from day 125 to day 240.

The order *Methanobacteriales* also exhibited a relatively even distribution across the three bioreactor samples (Fig. 4.6). However, in contrast to *Methanosaetaceae*, there was a decrease in 16S rRNA gene copy numbers from day 125 to 240, for example, R1 recorded a c.4.2-fold decrease from day 125 to 240. In R2, although there was an increase of c.1.3-fold recorded for *Methanobacteriales* from day 125 to day 128, there was a c.3.9-fold decrease between day 125 and 240 (Fig. 4.6).

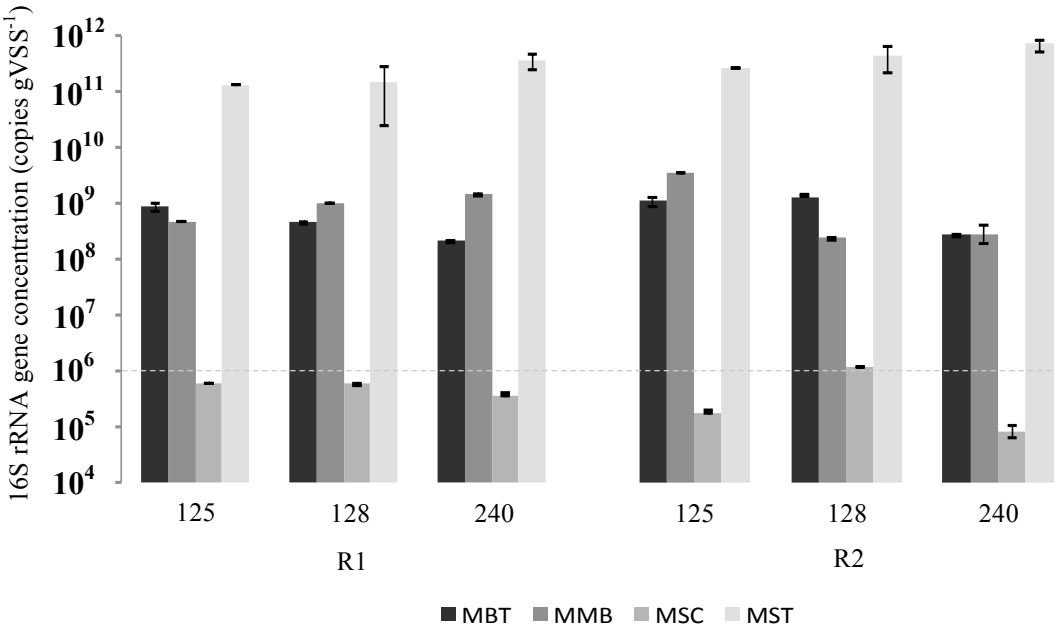


Figure 4.6 Absolute quantification of the 16S rRNA gene concentration of methanogens: **MBT** *Methanobacteriales*; **MMB** *Methanomicrobiales*; **MSC** *Methanosarcinicaeae*; **MST** *Methanosaetaceae* from biomass sampled on day 125 (37°C), 128 and 240 (15°C) for R1 and R2 bioreactors. Error bars indicate the standard deviation and are the result of two replicates. Dashed line relates to detection limit of assay (10^6)

The other hydrogenotrophic group, *Methanomicrobiales*, was also recorded at relatively consistent levels of 16S rRNA gene copies throughout the trial (Fig. 4.6). In R1, gene copy numbers ranged from 4.77 x 10^8 to 1.43 x 10^9 copies $gVSS^{-1}$ from day 125 to 240, representing a c. 3-fold increase. However, in contrast, R2 exhibited a >1-fold decrease from day 125 to 240 (Fig. 4.6).

The comparable 16S rRNA gene copy profiles for each group in each bioreactor suggests minimal change occurred in methanogenic community structure with respect to temperature. The prominence of *Methanosaetaceae* is evident with this group comprising >98% of the total measured methanogenic population in both bioreactors, i.e., the sum of all 16S rRNA gene concentrations quantified with the primer/probe sets used for this sample above the detection limit of 10^6 copies $gVSS^{-1}$. The other *Methanosarcinales* family measured, *Methanosarcinaceae*, was only detected above the level of quantification in this assay on day 128 in R2 (Fig. 4.6), with 1.04 x 10^6 copies g VSS^{-1} recorded. Thus, from this dataset it would appear that the *Methanosaetaceae* was the dominant acetoclastic methanogenic group active in these bioreactors.

4.3.6 DGGE

Comparative PCR-DGGE analysis of the archaeal community structure of R1 and R2 biomass on days 125 (37°C), 126, 128, 132 and 240 (15°C) indicated clear differences between DNA and cDNA profiles recorded (Fig. 4.7 [A]). The high number of cDNA 16S rRNA PCR fragments in comparison with DNA generated profiles was consistent in both R1 and R2 samples.

Although differences between DNA and cDNA profiles were clear, the DNA profiles demonstrated a grouping in UPGMA cluster analysis similar to that observed from cDNA profiles (Fig. 4.7 [B]). For example, in R1 there was 85% similarity for days 125, 126 and 128 in UPGMA analysis of DNA profiles, with > 90% similarity recorded for the corresponding cDNA profiles. This was also consistent in R2. Additionally, a temperature related distinction was observed in both

profiles, with the later sampling points in both R1 and R2 clustering together and displaying high similarity, for example, in R2 there was > 90% similarity recorded for day 132 and 240 DNA profiles, while > 85% similarity for corresponding cDNA profiles (Fig. 4.7 [B]).

The order *Methanosarcinales* was well represented by DGGE phylogenetic analysis, in agreement with the archaeal clone library analysis (Fig. 4.5; Fig. 4.8). Partial 16S rRNA sequences from six excised DGGE bands grouped with this order, with the family *Methanosaetaeae* evident once again. Three of these bands were prominent in DNA profiles, with a single band (10**;** Fig. 4.7[A]; Fig. 4.8) exhibiting high consistency between DNA and cDNA profiles in both R1 and R2. However, in contrast to clone libraries, the family *Methanosarcinaceae* was represented in DGGE analysis, with a single sequence from an excided cDNA band (18**;** Fig. 4.7[A]; Fig. 4.8) aligning with this group.

The order *Methanomicrobiales* was the most represented archaeal group in DGGE sequence analysis, with eight sequences aligning with this hydrogenotrophic group (Fig. 4.8). There was an even distribution between sequences originating from DNA and cDNA excised bands (Fig. 4.7[A]; Fig. 4.8), suggesting that, not only was this group prominent in R1 and R2 biomass during this trial, but had functional significance also. For the *Methanomicrobiales* order, four partial 16S rRNA sequences from excised cDNA bands grouped with *Methanolinea* (Fig. 4.7[A]; Fig. 4.8). Also, two sequences relating to cDNA bands were closely aligned to *Methanospirillum*, (Fig. 4.7[A]; Fig. 4.8), with this group not represented in archaeal clone library analysis (Fig. 4.5).

Finally, the phylum *Crenarchaeota* was also evident in the DGGE results with two sequences aligning closely with this group (Fig. 4.8). Interestingly, one sequence originated from a band that was prominent in both DNA and cDNA profiles for R1 and R2, whereas the other sequence was from a band was only evident in cDNA profiles (Fig. 4.7[A]; Fig. 4.8).

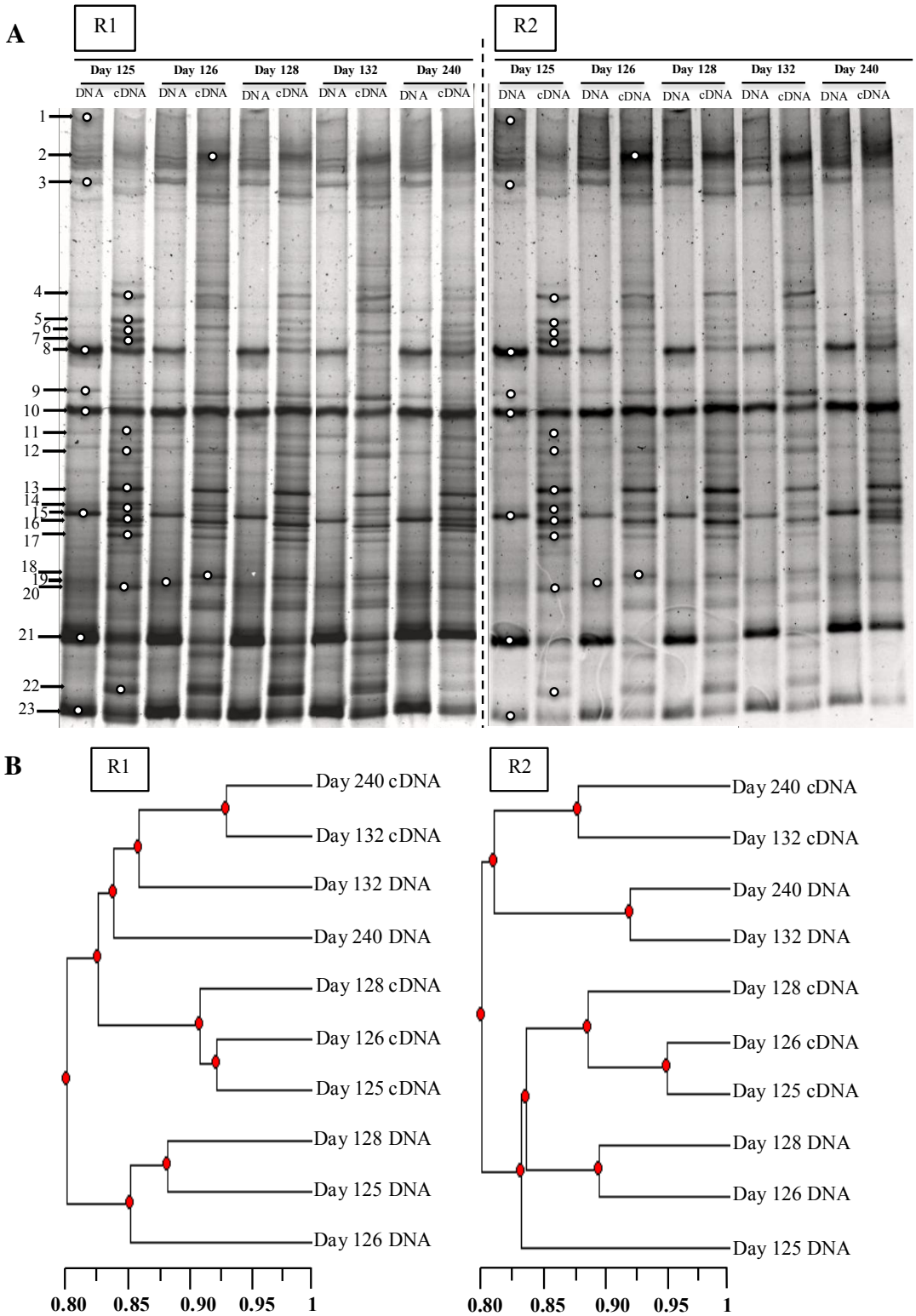


Figure 4.7 (**A**) DGGE analysis of archaea present in R1 and R2 at 37°C (Day 125), initial days after a temperature drop to 15°C (Days 126, 128 and 132) and also at the end of trial (Day 240). Arrows indicate DGGE bands for sequencing. (**B**) Archaeal UPGMA cluster analysis of 16S rRNA gene fragments generated from DGGE profiles. Similarity calculated by (Sorensins (Bray-Curtis) distance measurement

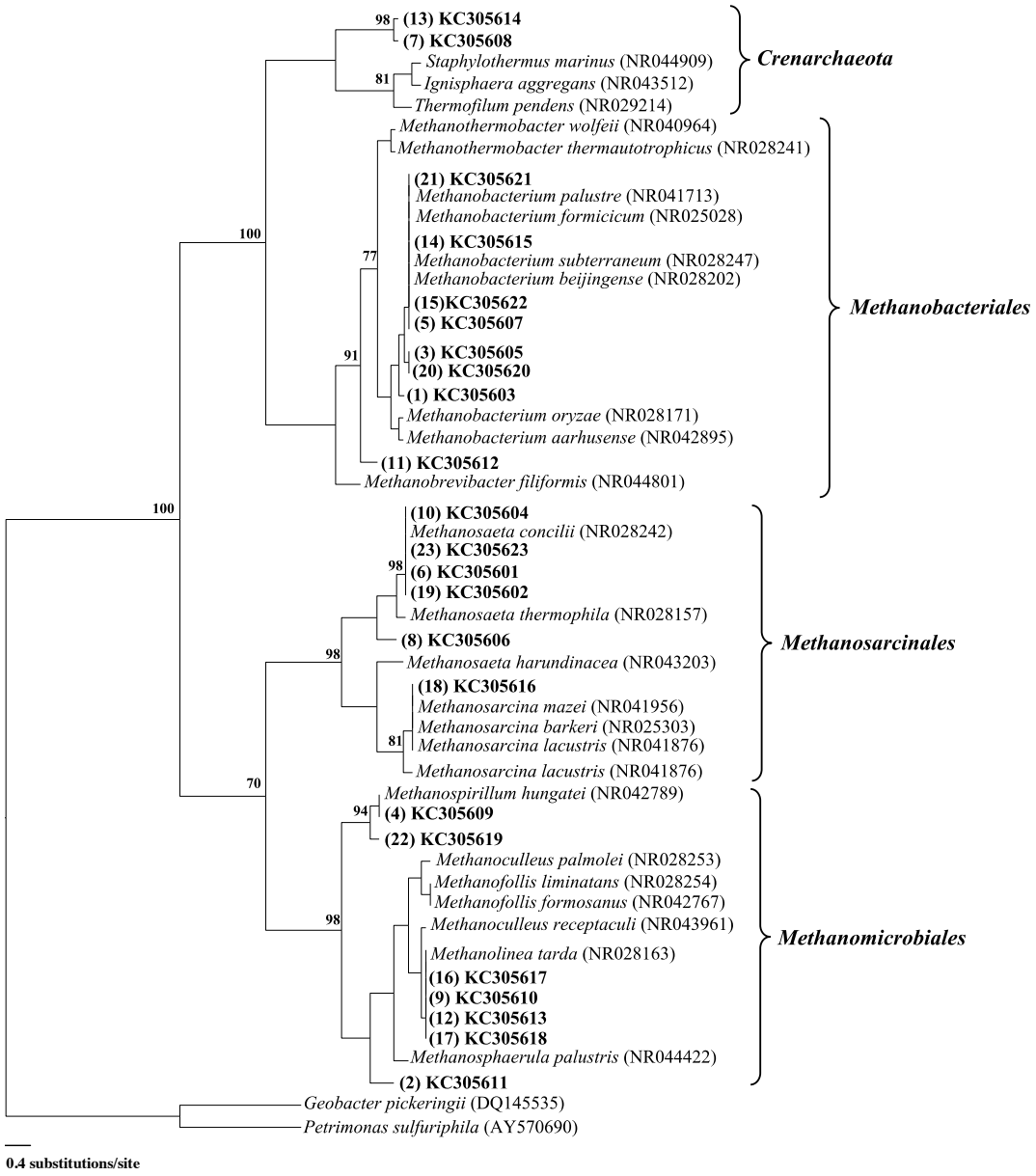


Figure 4.8 Phylogenetic affiliation of the archaeal 16S rRNA gene sequences obtained from DGGE analysis on day 125 (37°C), days 126, 128, 132 and 240 (15°C) bioreactor biomass, calculated using the GTR + gamma model DNA substitution implemented RAxML 7.0.3 (Stamatakis, 2006). Bootstrap replicates (total 1000 replicate samples) supporting the branching order are shown at relevant nodes. DGGE band relating to sequence is in parentheses.

4.3.7 MAR-FISH

MAR-positive cellular uptake of the radiolabelled substrates, [^{14}C]- acetate and H_2/CO_2, sodium-[^{14}C]-biocarbonate was recorded in both bioreactor biomass samples from days 125 (37°C) and 240 (15°C). Differences the in specific uptake patterns were apparent for both substrates at each sampling point, with differences also recorded between uptake profiles when both sampling points were compared (Table 4.5; Table 4.6).

Biomass sampled on day 125 exhibited efficient uptake of [^{14}C]- acetate, with cells positive for *Mx825* and *Sarc1551* probes comprising c.33% (R1) and c.46% (R2) of MAR-positive cells after 5 hrs of incubation at 37°C (Table 4.5). However, substrate uptake of radiolabelled acetate was slower at the same time-point when incubated at 15°C, with only c.8% (R1) and c.9% (R2) of MAR-positive cells relating to this group (Table 4.5). After 10 hrs of incubation at 37°C, the acetoclastic probes accounted for c.64% (R1) and c.68% (R2) of MAR-positive cells recorded, indicative of this groups significant functional importance at this sampling-point in both bioreactors (Table 4.5). MAR-positive cells relating to acetoclastic methanogens did not reach >50% until after 25 hours of incubation at 15°C. Biomass sampled from day 240 recorded an improvement in substrate uptake of acetate at both incubation temperatures, for example, after 5 hrs of incubation at 37°C c. 64% (R1) and c. 52% (R2) of MAR-positive cells related to acetoclastic methanogens (Table 4.6). Also, at 15°C there was an improvement recorded in comparison to day 125 biomass, with c. 54% (R1) and 64% (R2) of MAR-positive cells relating to acetoclastic probes after 15 hours of incubation (Table 4.6).

Hybridisation probes relating to the hydrogenotrophic group (*MB1174* and *MG1200b*) exhibited a reduced MAR-positive profile for day 125 biomass incubated with radiolabelled H_2, sodium bicarbonate when compared with acetoclastic substrate uptake (Table 4.5). At 37°C, this hydrogen utilising group did not reach > 50% of MAR-positive cells until after 15 hours of incubation, with < 50% of MAR-positive cells after 35 hours of incubation at 15°C. However, biomass from day 240

showed an increase in hydrogenotrophic activity with c.57% (R1) and 63% (R2) of MAR-positive cells relating to hydrogenotrophic probes after 15 hrs of incubation at 15°C (Table 4.6). A similar result was recorded when biomass was incubated at 37°C for both R1 and R2. Selected images of MAR-positive cells from radiolabelled substrate incubations in coordination with fluorescent probe images are shown in Fig. 4.9. Although it was evident that *Methanosaeta* (*Mx825*) and *Methanosarcina* (*Sarc1551*) were the prominent functional groups in acetate utilisation (Fig. 4.9 [A, B]), it was observed that hydrogenotrophic functional groups other than *Methanobacteriales* (*MB1174*) and *Methanomicrobiales* (*MG1200b*) were also utilising this substrate during these incubations, especially in biomass sampled from day 125 (Fig. 4.9 [c]). Comparative MAR-FISH substrate uptake images after 15 hrs of incubation comparing the two biomass samples is shown in Appendix 3.

The corrected total cell fluorescence (CTCF) measurements recorded with acetoclastic and hydrogenotrophic probes showed consistency with the substrate uptake profiles from biomass sampled on days 125 and 240 from R1 and R2. Specifically, the relative fluorescence units (RFUs) calculated showed an increase for both functional groups in relation to sampling point (Fig. 4.10). However, there was a more defined increase recorded for the hydrogenotrophic group, with a c.1.7-fold (R1) and c.2.1-fold (R2) increase in RFUs from day 125 to day 240, which was reflected in the substrate uptake activity also (Table 4.5; Table 4.6).

Table 4.5 Relative abundance of MAR-positive cells following uptake of ^{14}C-labelled substrates in coordination with archaeal hybridisation probes. Results are from biomass sampled on day 125 (37°C).

Biomass	Day 125															
Temperature	37°C								15°C							
Substrate	^{14}C Acetic Acid				H_2, Sodium ^{14}C bicarbonate				^{14}C Acetic Acid				H_2, Sodium ^{14}C bicarbonate			
Time (hours)	5	10	15	20	5	10	15	20	5	15	25	35	5	15	25	35
Mx825 + *SarcI551* (R1)	++	+++	+++	+++	–	+	+	++	+	+	++	+++	–	–	+	++
Mx825 + *SarcI551* (R2)	++	+++	+++	+++	–	+	++	++	+	+	++	+++	–	–	+	++
MB1174 + *MG1200b* (R1)	–	–	–	–	+	++	++	+++	–	–	–	–	–	+	++	++
MB1174 + *MG1200b* (R2)	–	–	–	–	+	++	++	+++	–	–	–	–	–	+	++	++

Notes:
– : No uptake of substrate recorded
+ : < 10% of the total MAR-positive cells
++ : 10-50% of the MAR-positive cells
+++ : > 50% of the MAR-posotive cells

Table 4.6 Relative abundance of MAR-positive cells following uptake of ^{14}C-labelled substrates in coordination with archaeal hybridisation probes. Results are from biomass sampled on day 240 (15°C).

Biomass	Day 240															
Temperature	37°C								15°C							
Substrate	^{14}C Acetic Acid				H_2, Sodium ^{14}C bicarbonate				^{14}C Acetic Acid				H_2, Sodium ^{14}C bicarbonate			
Time (hours)	5	10	15	20	5	10	15	20	5	15	25	35	5	15	25	35
Mx825 + *SarcI551* (R1)	+++	+++	+++	+++	–	++	++	++	++	+++	+++	+++	+	+	++	++
Mx825 + *SarcI551* (R2)	+++	+++	+++	+++	–	+	++	++	++	+++	+++	+++	+	++	++	++
MB1174 + *MG1200b* (R1)	–	–	–	–	++	+++	+++	+++	–	–	–	–	++	+++	+++	+++
MB1174 + *MG1200b* (R2)	–	–	–	–	++	+++	+++	+++	–	–	–	–	++	+++	+++	+++

Notes:
– : No uptake of substrate recorded
+ : < 10% of the total MAR-positive cells
++ : 10-50% of the MAR-positive cells
+++ : > 50% of the MAR-posotive cells

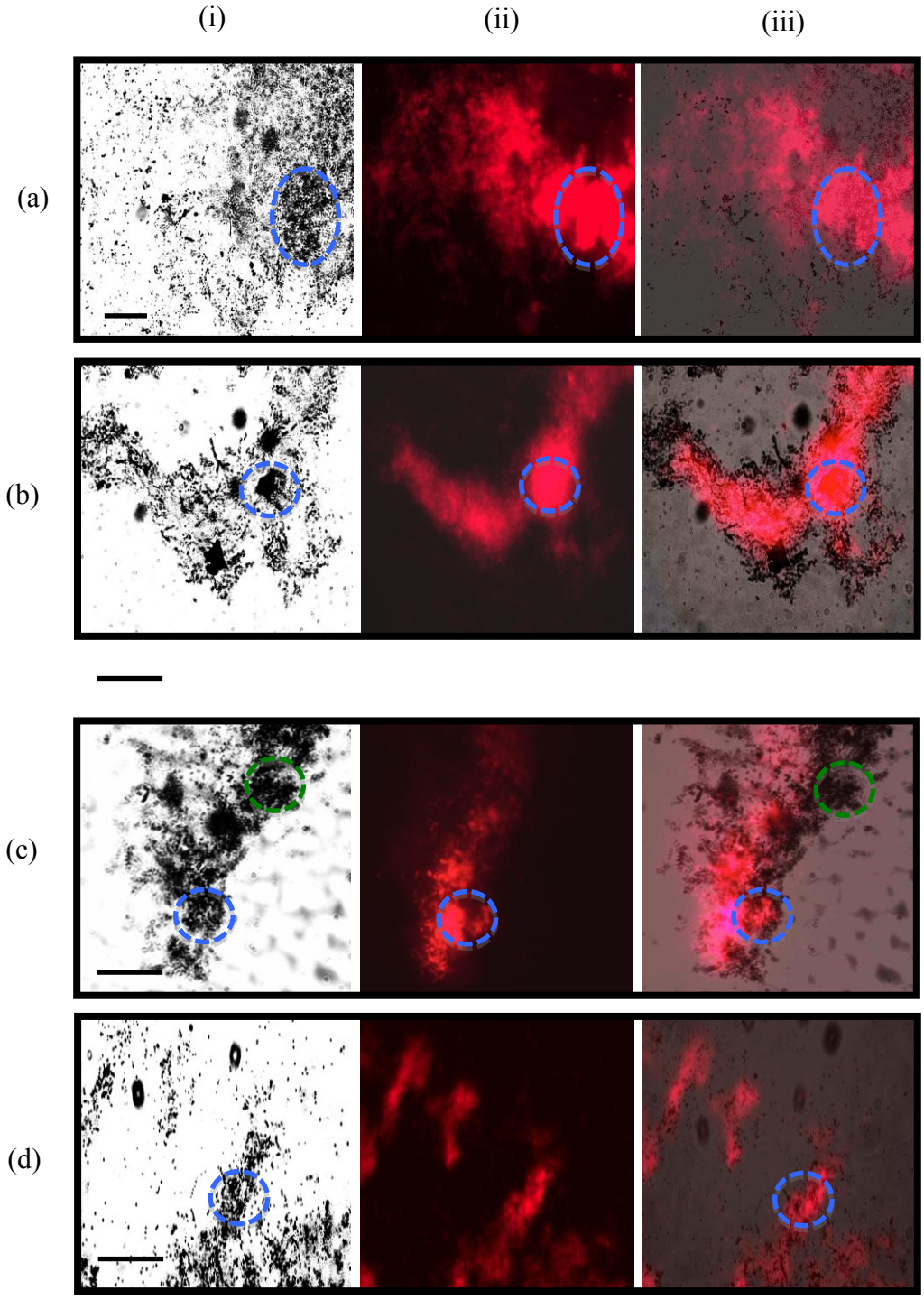


Figure 4.9 Selected MAR-FISH images from [^{14}C] acetic acid incubations from days (a) 125 and (b) 240 and from H_2, sodium-[^{14}C]-bicarbonate incubations from days (c) 125 and (d) 240. For each sample, (i) MAR (silver grains), (ii) FISH signal and (iii) MAR-FISH overlay are shown, depicting MAR-positive cells (indicated by dashed blue line) and other areas of substrates uptake (dashed green line). Scale bar is 10 μm.

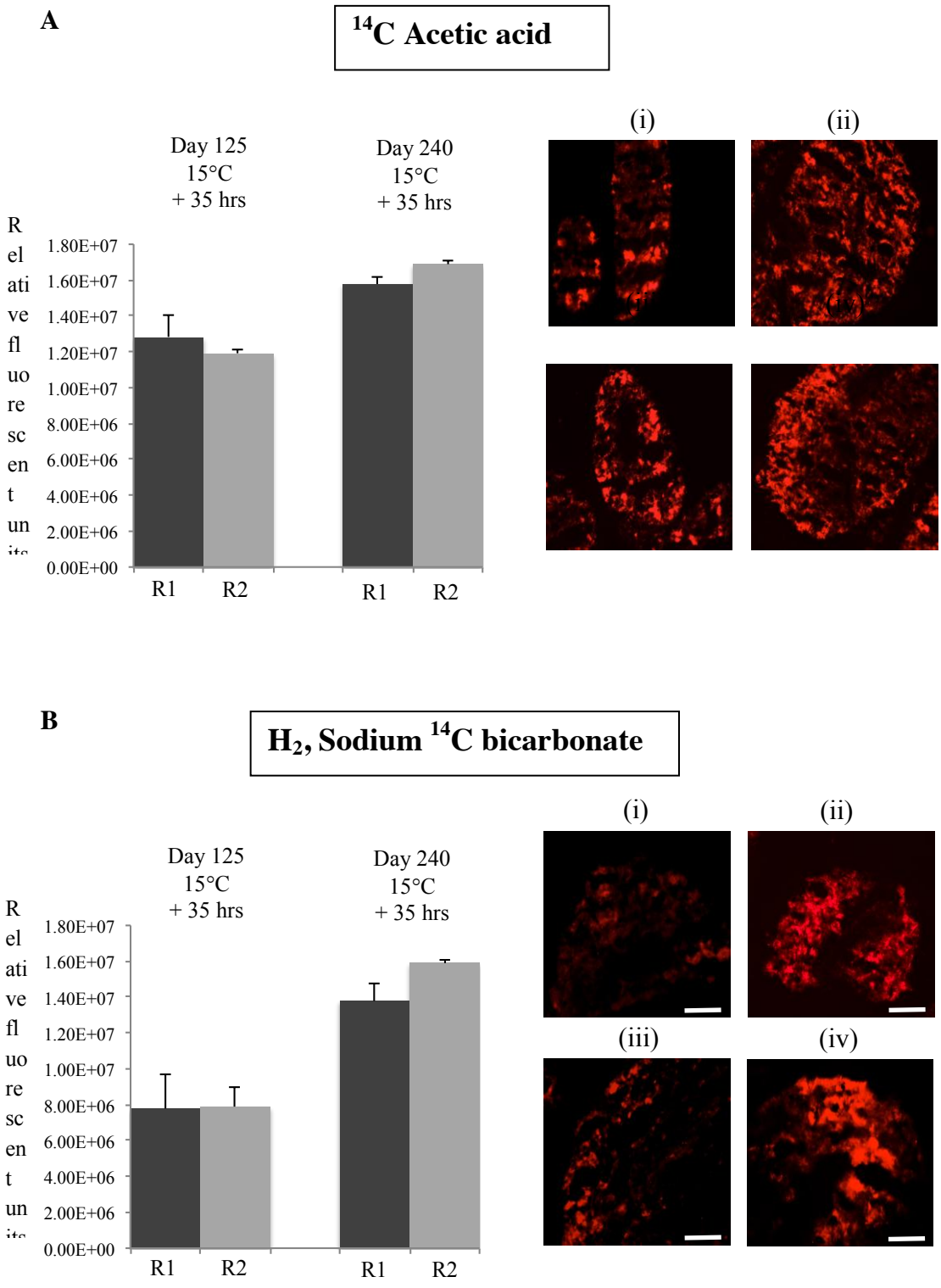


Figure 4.10 Relative fluorescence of (A) acetoclastic archaea (*Mx825* + *SarcI551*) and (B) hydrogenotrophic acrchaea (*MB1174* + *MG1200b*) in granular sections on days 125 and 240 after incubation at 15°C for 35 hours. Scale bar is 100 μm.

4.3.8 Metaproteomics

Positive identification of fifty-nine proteins through nLC-ESI MS/MS analysis was recorded (Table 4.7; Table 4.8). These proteins were excised from distinct 2D-gel protein spots run with samples taken on days 125 (37°C), 128 and 240 (15°C) of trial from R1 and R2.

A total of seventeen proteins were associated with bacterial species, fourteen of which were members of the phylum *Proteobacteria*, with the *Actinobacteria*, *Firmicutes* and *Thermatagoe* represented by a single protein each (Table 4.7). The assigned functions of these proteins originating from bacterial groups comprised of two main categories, the first being general metabolism and biosynthesis (Fig 4.11). Proteins associated with the metabolism of wastewater components such as acetate (acetate CoA synthetase[ACS]) and ethanol (iron-containing alcohol dehydrogenase) were identified in all samples, and were affiliated with the *Proteobacteria* group (Table 4.7). Interestingly, acetate CoA synthetase (Syn_02635) and alcohol dehydrogenase (Despr_2940) proteins were upregulated in the low-temperature biomass when compared with 37°C sample on day 125 (Table 4.7). Also, a catalase protein (SacctDRAFT_224), which is important in amino acid metabolism was identified and associated with *Actinobacteria*.

Transporters and other membrane-associated proteins were also recorded, with the *Firmicutes* and *Thermatogoe* represented by this functional category along with *Proteobacteria*. These included two extracellular ligand-binding proteins associated with beta- (Sfum_3105) and delta- (Varpa_2832), which were conserved in both bioreactors throughout the trial. A coenzyme A transferase protein (DealDRAFT_0496), assigned to a *Syntrophus* species was upregulated in low-temperature biomass (Table 4.7). Two molecular chaperones were also evident, which included a DnaK protein (Sala_2058) expressed at a higher level at 37°C when compared with both low-temperature samples, with the reverse recorded for the other chaperone protein (Syn_01983; Table 4.7).

Species assigned affiliated forty-two proteins with archaeal representatives, relating to the orders *Methanobacteriales* and *Methanomicrobiales*, the family *Methanosaetaceae*, and also the Phylum *Crenarchaeota* (Table 4.8). Methanogenesis was the most prominent functional category of proteins originating from archaea (Fig. 4.11), with proteins involved in acetate (CO dehydrogenase/acetyl-CoA synthase) and CO_2 (F_{420}-dependent methylenetetrahydromethanopterin reductase) methanogenesis identified in each biomass sample relating to the three euryarchaeal groups listed above (Table 4.8). Following on from bacterial results, proteins identified relating to acetate metabolism in archaea were found to be upregulated in low-temperature biomass, which included the protein responsible for acetyl A catabolic process (Mcon_1326), assigned to a *Methanosaeta* species. Interestingly, two oxidoreductase proteins with similar function and associated with methanogenesis from CO_2 exhibited different results. These two proteins were assigned to *Methanobacterium* species, with the first protein (Mswan_1889) expressed at a higher level at 37°C compared with the 15°C sample on day 240, with the reverse recorded for the second protein (Mtbma_c03270; Table 4.8).

Also identified was a hydrogenase protein associated with CO_2 and formate metabolism (Mcon_2320). Surprisingly, this protein was assigned to the acetoclastic *Methanosaeta*, and was once again expressed at a higher level in day low-temperature biomass when compared with the 37°C sample on day 125 (Table 4.8). A protein responsible for catalysing the final setup in methane formation, methyl coenzyme M reductase (MCR), was also represented. There were several MCR proteins identified with differential expression profiles recorded as a function of temperature (Table 4.8).

Another functional category well represented by proteins associated with the archaeal group was metabolislm and biosynthesis (Fig. 4.11). The majority of these were assigned to *Methanosaeta* species, including proteins associated with purine (Mcon_0554) and lysine (Mcon_0799) biosynthesis. A number of these proteins were expressed at a higher level in low-temperature biomass, which included the

bifunctional enzyme (Mcon_1383), which can function in the ribulose monophosphate pathway as well as formaldehyde activation with a possible role in methanogenesis (Table 4.8). Also identified was a fructose-binding protein (Mhun_1036) assigned to *Methanospirillum* species, which was prominent in low-temperature biomass. This hydrogenotrophic group was well represented in this dataset with seven proteins identified, which were primarily upregulated in the sub-mesophilic biomass samples (Table 4.8).

Also recorded were proteins associated with substrate transportation, energy generation, DNA binding and repair, and a surface secreted protein associated with the S-layer of *Methanosaeta* (Mcon_1153). Interestingly, there were no proteins relating to molecular chaperones or stress response related to archaea recorded in this dataset, although a protein associated with DNA binding and repair (Saci_0192) was identified, which was assigned to a *Sulfolobus* species, a member or the order *Sulfolobales* in the phylum *Crenarchaeota*. This protein was upregulated in 37°C when compared with both 15°C samples (Table 4.8).

Table 4.7 Identification and putative function assigned to bacterial proteins excised from 2D-gels derived from bioreactor biomass. Ratios relate to differential abundance of protein for the three biomass samples day 125 (A), day 128 (B) and day 240 (C). Significantly expressed proteins are in bold.

Gene locus	Protein name	% Cov	Assigned species	Differential abundance ratios					
				A/B		A/C		B/C	
				R1	R2	R1	R2	R1	R2
Syn_02635	Acetyl-CoA synthetase	9	*Syntrophus aciditrophicus*	0.74	**2.88**	**0.29**	0.94	**0.37**	**0.32**
Sfum_1047	Adenylylsulfate reductase β-subunit	17	*Syntrophobacter fumaroxidans*	1.44	1.51	0.93	0.58	1.39	1.32
SaccyDRAFT_2244	Catalase	15	*Saccharomonospora cyanea*	1.49	**4.12**	**0.37**	**0.38**	**0.25**	**0.09**
Gura_2389	Iron-containing alcohol dehydrogenase	19	*Geobacter uraniireducans*	0.98	**2.63**	**0.35**	0.81	**0.38**	**0.30**
Despr_1637	Iron-containing alcohol dehydrogenase	36	*Desulfobulbus propionicus*	1.15	**2.83**	**0.16**	**0.36**	**0.34**	**0.12**
Despr_1373	*Sulfate adenylyltransferase*	43	*Desulfobulbus propionicus*	1.05	1.55	**0.36**	**0.36**	**0.34**	**0.23**
Despr_2940	Dissimilatory sulfite reductase α-subunit	9	*Desulfobulbus propionicus*	0.84	0.79	**3.81**	**2.21**	**4.52**	**4.09**
Sfum_1703	Succinyl-CoA synthetase β-subunit	15	*Syntrophobacter fumaroxidans*	0.69	2.79	**0.16**	**0.38**	**0.23**	**0.14**
Sfum_3105	Extracellular ligand-binding receptor	47	*Syntrophobacter fumaroxidans*	0.59	1.65	0.69	0.57	1.16	0.79
Varpa_2832	Extracellular ligand-binding receptor	23	*Variovorax paradoxus*	1.89	1.28	1.29	1.53	1.95	1.89
VTlet_1531	Extracellular solute-binding protein	8	*Thermotoga lettingae*	**2.71**	**2.27**	**2.93**	**2.62**	1.42	1.21

Gene locus	Protein name	% Cov	Assigned species	Differential abundance ratios					
				A/B		A/C		B/C	
				R1	R2	R1	R2	R1	R2
Despr_3315	PAAT family amino acid ABC transporter substrate-binding protein	14	*Desulfobulbus propionicus*	0.68	1.78	**0.21**	**0.38**	**0.22**	**0.34**
Varpa_2489	Phosphate ABC transporter periplasmic phosphate-binding protein	18	*Variovorax paradoxus*	**3.22**	**2.71**	**8.26**	**4.87**	**2.52**	**3.24**
DealDRAFT_0496	Coenzyme A transferase	13	*Dethiobacter alkaliphilus*	1.03	**2.67**	**0.35**	0.97	**0.35**	**0.37**
Syn_01983	Chaperone protein	45	*Syntrophus aciditrophicus*	0.60	**2.09**	**0.22**	1.96	**0.37**	0.94
Sala_2058	Molecular chaperone DnaK	7	*Sphingopyxis alaskensis*	1.86	1.49	**2.16**	**3.87**	**2.56**	**2.26**
Despr_3276	Phage tail sheath protein	8	*Desulfobulbus propionicus*	**2.16**	**2.42**	**2.86**	**3.47**	**2.56**	**2.26**

Table 4.8 Identification and putative function assigned to archaeal proteins excised from 2D-gels derived from bioreactor biomass. Ratios relate to differential abundance of protein for the three biomass samples day 125 (A), day 128 (B) and day 240 (C). Significantly expressed proteins are in bold.

Gene locus	Protein name	% Cov	Assigned species	Differential abundance ratios					
				A/B		A/C		B/C	
				R1	R2	R1	R2	R1	R2
Mcon_1330	CO dehydrogenase/acetyl-CoA synthase complex β-subunit	23	*Methanosaeta concilii*	**4.86**	**2.32**	**2.86**	1.47	0.56	**0.26**
Mcon_1326	CO dehydrogenase/acetyl-CoA synthase complex δ-subunit	30	*Methanosaeta concilii*	0.83	1.33	**0.27**	**0.34**	**0.32**	**0.26**
Mtbma_c03270	F_{420}-dependent methylenetetrahydromethanopterin reductase	65	*Methanothermobacter marburgensis*	1.22	0.98	**0.37**	**0.36**	**0.30**	**0.36**
Mswan_1889	5,10-methylenetetrahydromethanopterin reductase	76	*Methanobacterium sp*.	1.32	1.56	**3.51**	1.53	**2.63**	**2.97**
Mcon_2320	Coenzyme F_{420} hydrogenase β-subunit	29	*Methanosaeta concilii*	**3.64**	**2.68**	**0.38**	**3.20**	**0.10**	1.19
Mcon_0759	Methyl-coenzyme M reductase α-subunit	20	*Methanosaeta concilii*	**3.68**	**9.70**	**5.88**	**2.72**	1.59	**0.28**
Mswan_2056	Methyl-coenzyme M reductase α-subunit	42	*Methanobacterium subterraneum*	**2.98**	0.69	**21.34**	0.87	**7.15**	1.26
Mswan_2060	Methyl-coenzyme M reductase β-subunit	57	*Methanobacterium sp*.	**2.30**	**2.20**	**11.59**	**2.46**	**5.02**	1.11

Gene locus	Protein name	% Cov	Assigned species	Differential abundance ratios					
				A/B		A/C		B/C	
				R1	R2	R1	R2	R1	R2
Mcon_0762	Methyl-coenzyme M reductase β-subunit	31	*Methanosaeta concilii*	1.92	**2.57**	**0.36**	**2.35**	**0.18**	0.91
Mhun_2144	Methyl-coenzyme M reductase β-subunit	40	*Methanopsirillum hungatei*	**3.39**	**2.90**	**0.37**	**0.43**	**0.10**	**0.32**
Mhar_0495	Methyl-coenzyme M reductase β-subunit	33	*Methanosaeta harundinacea*	**10.89**	**2.20**	**5.51**	1.29	**0.50**	0.59
Mhun_2147	Methyl-coenzyme M reductase γ-subunit	48	*Methanopsirillum hungatei*	**2.46**	1.68	**2.22**	0.88	0.90	0.51
Mth_1165	Methyl-coenzyme M reductase γ-subunit	30	*Methanothermobacter thermautotrophicus*	1.24	0.91	**0.34**	**0.36**	**0.30**	**0.36**
Mhun_2148	Coenzyme-B sulfoethylthiotransferase	18	*Methanopsirillum hungatei*	0.90	**10.77**	**0.33**	**3.10**	**0.22**	**0.28**
Mcon_2324	Phosphate binding protein	27	*Methanosaeta concilii*	1.15	1.22	**2.42**	**2.11**	**2.08**	1.75
Mcon_1153	S-layer-related duplication domain-containing protein	50	*Methanosaeta concilii*	1.27	0.51	0.82	0.63	0.64	1.22
Mcon_1628	Periplasmic binding protein	19	*Methanosaeta concilii*	1.49	0.85	1.91	1.88	1.28	1.87
Mcon_0477	Periplasmic binding protein	46	*Methanosaeta concilii*	1.56	1.85	0.79	**2.36**	0.51	1.27
Mcon_2202	Periplasmic binding protein	24	*Methanosaeta concilii*	1.84	1.15	1.59	1.38	0.61	1.71
Mcon_0484	ABC transporter substrate-binding protein	25	*Methanosaeta concilii*	**2.87**	**4.35**	**4.69**	**3.27**	**2.69**	**3.71**
Mcom_0157	Enolase	63	*Methanosaeta concilii*	**2.33**	**1.88**	**13.05**	**2.19**	**5.44**	**1.16**

Gene locus	Protein name	% Cov	Assigned species	Differential abundance ratios					
				A/B		A/C		B/C	
				R1	R2	R1	R2	R1	R2
Mcon_2515	V-type ATP synthase β-subunit	5	*Methanosaeta concilii*	0.84	**0.29**	**3.81**	1.21	**4.51**	**4.09**
Mcon_1383	Bifunctional enzyme fae/hps	27	*Methanosaeta concilii*	1.45	1.78	**0.38**	**0.26**	**0.36**	**0.31**
Mcon_0165	TPR repeat-containing protein	29	*Methanosaeta concilii*	0.16	2.02	**0.06**	**0.36**	**0.18**	**0.40**
Mcon_1906	Manganese-dependent inorganic pyrophosphatase	47	*Methanosaeta concilii*	1.51	**14.32**	**0.36**	**2.32**	**0.23**	**0.16**
Mcon_0035	Beta-lactamase domain-containing protein	25	*Methanosaeta concilii*	0.90	**3.88**	**0.32**	**0.29**	**0.35**	**0.07**
Mcon_0429	3-hexulose-6-phosphate synthase	16	*Methanosaeta concilii*	0.98	**1**.24	**0.42**	**0.31**	**0.38**	**0.37**
Mcon_2723	Peptidase, ArgE/DapE family	23	*Methanosaeta concilii*	0.73	**1**.89	**0.22**	**0.43**	**0.44**	**0.34**
Mcon_0554	Adenylosuccinate synthetase	16	*Methanosaeta concilii*	**3.18**	**2.45**	**12.87**	**6.12**	**4.14**	**3.26**
Mcon_2418	Aminotransferase	13	*Methanosaeta concilii*	0.93	**2.71**	**0.19**	0.62	**0.20**	**0.23**
Mcon_0799	Homoserine dehydrogenase	6	*Methanosaeta concilii*	**3.27**	**2.79**	**4.59**	**3.29**	1.47	1.78
Mcon_0412	Ketol-acid reductoisomerase	24	*Methanosaeta concilii*	0.77	1.96	1.41	1.69	1.81	0.91
Mcon_2555	Bifunctional short chain isoprenyl diphosphate synthase	38	*Methanosaeta concilii*	1.76	1.25	1.87	1.63	0.82	1.19

Gene locus	Protein name	% Cov	Assigned species	Differential abundance ratios					
				A/B		A/C		B/C	
				R1	R2	R1	R2	R1	R2
Mcon_0804	Dihydrodipicolinate synthase	18	*Methanosaeta concilii*	**0.50**	**2.98**	**0.12**	**0.25**	**0.23**	**0.18**
Mhun_1036	Fructose-bisphosphate aldolase	13	*Methanosirillium hungatei*	0.74	**2.72**	**0.19**	1.01	**0.25**	**0.37**
Saci_0192	XPD/Rad3 related DNA helicase	70	*Sulfolobus acidocaldarius*	**3.86**	**9.27**	**5.88**	**2.72**	1.59	1.26
Mcon_2173	Hypothetical protein	13	*Methanosaeta concilii*	1.31	**2.41**	**0.35**	1.38	**0.26**	0.57
Mcon_2188	Hypothetical protein	34	*Methanosaeta concilii*	0.77	0.60	**5.36**	**2.67**	**7.25**	**2.57**
Mhun_0996	Hypothetical protein	12	*Methanosirillium hungatei*	0.97	0.98	1.24	0.97	1.64	1.93
Mhun_2513	Hypothetical protein	36	*Methanosirillium hungatei*	**3.25**	**5.98**	1.17	1.39	**0.36**	**0.23**
Mhun_2263	Hypothetical protein	38	*Methanosirillium hungatei*	**3.29**	**3.96**	1.12	1.40	**0.35**	**0.35**
MettaDRAFT_0691	Hypothetical protein	21	*Methanolinea tarda*	**4.23**	**2.87**	1.92	1.30	**0.25**	**0.45**

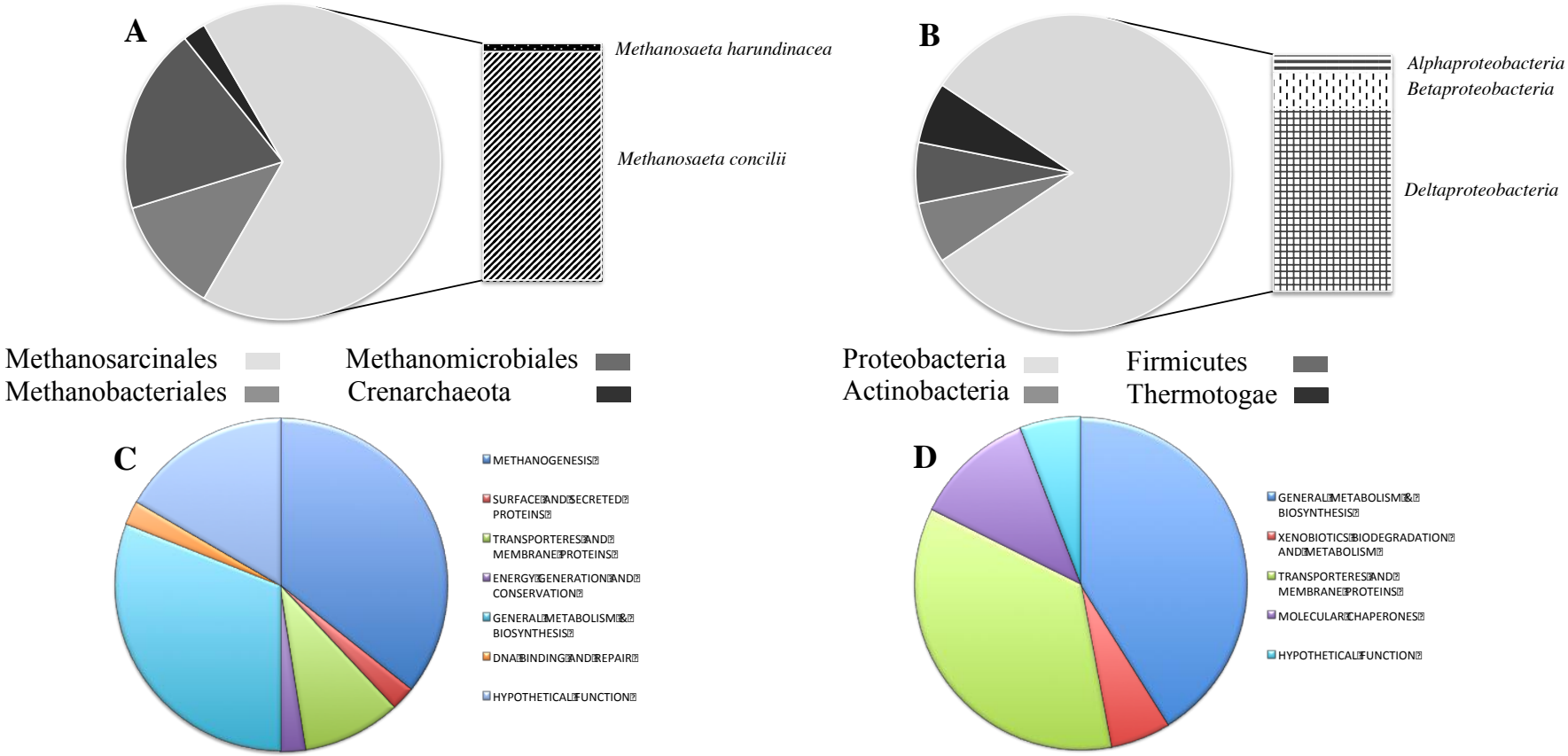


Figure 4.11 Illustrated representation of metaproteomics results with (A) proteins identified assigned to archaeal species and (B) proteins identified assigned to bacterial species. The functional categories of proteins identified for (C) the archaeal and (D) bacterial species are also presented

4.4 Discussion

This study demonstrated the feasability of employing PCR-based and PCR-independent techniques in a polyphasic approach for the study of microbial community structure, while also determining key functional groups associated with successful LTAD operation. Both bioreactors exhibited high treatment efficiencies, recording >80% COD removal for the majority of trial, when operated at 15°C. This was in conjunction with high CH_4 levels in biogas, with mean CH_4% rarely falling below 50%, regardless of operating temperature. High levels of efficient COD removal and CH_4 production have been recorded in many previous LTAD laboratory-based studies treating domestic (Cui *et al*., 2007), acidified (Kato *et al*., 1994; Lettinga *et al*., 1999), brewery (Connaughton *et al*., 2006), food-processing (Viraraghavan & Kikkeri, 1990) and pharmaceutical (Enright *et al*., 2005) wastewaters. By employing a relatively simple VFA-based wastewater with a HRT of 24 hrs, the primarily focus of this study could be focused on the microbial dynamics underpinning this LTAD system.

Clone library analysis revealed *Proteobacteria* to be predominant in mesophilic biomass, comprising > 45% of bacterial community in both bioreactors on day 125. Previous studies have found this phylum to be important in LTAD, with Enright *et al*. (2007), suggesting that this group in association with *Firmicutes* and green sulphur-like bacteria were responsible for the low-temperature treatment of a solvent-based wastewater. An increased distribution in *Protreobacteria*-like species was apparent in biomass sampled on day 240, which included a *Syntrophus*-like clone comprising > 10% of this bacterial clone library. Syntrophic bacteria have been characterised as pairing with methanogenic groups, for example, in the degradation of organic compounds such as benzoate (Schocke & Schink, 1997) alkanes in crude oil (Gray *et al*., 2011) propionate (de Bok *et al*., 2004) and acetate (Hattori *et al*., 2005). The degradation of organic compounds in anaerobic conditions requires that low hydrogen partial pressure is maintained, which allows the reaction to yield the minimum amount of energy required for the acetate-oxidising partner (McInerney *et al*., 2007). The finding of an aceteate activation enzyme assigned to a *Syntrophus* species (Syn_02635), which was upregulated in sub-mesophilic biomass

(Table 4.7) would suggest that this group was functionally important, with a possible role in acetate and/or propionate oxidation in sub-mesophilic biomass. The other bacterial groups represented in clone libraries showed a level of consistency between sampling points, with *Synergistetes* increasing from 6-14% of bacterial community from day 125 to 240 in R2, representing the only considerable increase recorded. The metaproteomics results also highlighted the dominance of *Proteobacteria* in these biomass samples, comprising 14 (82%) of the 17 proteins assigned to bacterial species. The proteins identified related primarily to the metabolism and cellular transport functional categories (Fig. 4.11). These included two ethanol metabolism proteins relating to separate *δ-proteobacteria* species, *Geobacter* (Gura_2389) and *Desulfobulbus* (Despr_1637). Interestingly, a protein relating to a *Thermotogae* species was also identified, with this group not represented in clone library analysis.

Archaeal community profiles exhibited interesting findings through the different approaches employed. Clone library analysis recorded 17 OTUs, with *Methanosaeta*-like organisms predominant at both sampling points (Fig. 4.5). This finding, in coordination with low acetate VFA measurements throughout the trial (R1 [mean 0.13 mg COD^{-1}], R2 [mean 0.15 $mgCOD^{-1}$]), suggest that *Methanosaetacea* were the primary methanogenic functional group underpinning both bioreactors. This was in conjunction with metaproteomic data, which affiliated 28 (66%) proteins to this acetoclastic family. In addition to proteins relating to acetate metabolim (Mcon_1330; Mcon_1326), a coenzyme F_{420} hydogenase protein (Mcon_2320) was also assigned to this group. This enzyme is usually associated with methylotrophic methanogenesis, which requires reduced F_{420} (F_{420}, H_2; Hendrickson & Leigh., [2008]). How *Methanosaetaceae* produces sufficient energy via acetate metabolism has been an issue of debate due to this group employing an acetyl CoA synthetase (ACS) system, requiring two ATP molecules (Smith & Ingrim-Smith, 2007). One suggestion is that F_{420}-reducing hydrogenase may be involved in membrane-bound ferredoxin:heterodisulfide oxidoreductase system for energy conservation distinct from those well-studied in *Methanosarcina* species (Welte & Deppenmeier, 2011). Although other proteins in this proposed complex were not identified, it is compelling that direct evidence of energy maintenance

strategies of *Methanosaetaceae* is presented in this set of results, with this hydrogenase protein also found to be upregulated in low-temperature biomass in both bioreactors (Table 4.8). Other proteins assigned to *Methanosaeta* species and expressed at a higher level in low-temperature biomass included the bifunctional formaldehyde activating enzyme (Mcon_1383), which as discussed in chapter 2, relates to methylotrophic bacterial metabolism where formaldehyde is formed as an intermediate of the pathway (Chistoserdova *et al*, 1998; Groenrich *et al*, 2005), but may also have a function in 'reverse' ribulose monophospahte pathway (RuMP) (Goenrich *et al*, 2005, Soderberg, 2005). SMA profiles also gave an indication of the importance acetoclastic methanogenesis was in these bioreactors, with activity profiles higher than hydrogenotrophic activity in both bioreactors at the end of the trial (Table 4.2). Also, qPCR rsults found *Methanosaetacea* to comprise ~ 98% of total measured methanogenic population in both bioreactors (Fig. 4.6). MAR-FISH profiles confirmed that acetoclastic methanogenesis was occurring significantly at both operation temperatures, with > 50% of MAR-positive cells relating to this acetoclastic methanogenesis after 15 hours of incubation (Table 4.5).

Although clear from results that *Methanocaetaeae* were actively functioning in both LTAD bioreactors, clones assigned to this group did show a reduction from day 125 (R1 [82%], R2 [62%]) to day 240 (R1 [68%], R2 [51%]). Though not clear through clone library and qPCR analysis, the community profiles generated from DGGE analysis in coordination with proteomics gave clear insights into other methanogenic groups functioning during this trial. For instance, a DGGE band associated with another acetoclastic methanogen, *Methanosarcina,* appeared in cDNA profiles during days immediately after temperature drop to 15°C (Fig. 4.7[A]), with this genera not detected in any other method applied. As this band was absent on days 125 and 240 it is possible that due to thermodynamic impacts on dominant organisms (e.g. *Methanosaeta*) brought about by the decrease in operation temperature, this methanogenic group was able to proliferate for a short period of time. This was in correlation with recorded acetate accumulation in both bioreactors during this time period, previously found to be associated with *Methanorsarcina* detection (Fig. 4.2; Enright *et al*., 2007).

Although *Methanomicrobiales* comprised only 10% (R1) and 12% (R2) of archaeal clones on day 240, this order accounted for 30% of partial 16S rRNA gene sequences obtained from DGGE band analysis. Also, two DGGE sequences closely aligned with a *Methanospirillum* species were obtained from bands more prominent in cDNA bioreactor profiles, suggesting an important functional role in sub-mesophilic biomass, for example, homoacetogenic activity with *Syntrophus* species (Gray *et al*., 2011). The functional importance of *Methanospirillum* during the trial was confirmed through the identification of seven proteins varying in function assigned to this genus, with the majority of these proteins upregulated in low-temperature biomass (Table 4.8).

The order *Methanobacteriales* had eight DGGE sequences assigned to this group, obtained from bands prevalant on both DNA and cDNA profiles (Fig. 4.7[A]). There appeared to be a succession within this order in relation to temperature and function, with the protein responsible for methane formation from methyl coenzyme-M assigned to four different *Methanobacteriales* species. Two of these were upregulated in low-temperature biomass in both bioreactors (Mtbma_c03270; Mth_1165). Higher hydrogenotrophic SMA profiles on day 240 compared with 125 would suggest that increased hydrogenotrophic methanogenic activity was apparent at 15°C compared with 37°C. It is possible that the temperature decrease applied a selective pressure facilitating the growth and function of hydrogenotrophic methanogens. MAR-FISH profiles suggested this also with > 50% of MAR-positive cells assigned to hydrogenotrophic methanogens after 15 hours on day 240, compared with < 50% MAR-positive cells after 35 hours of incubation on day 125. Also, there was a > 1.5-fold increase in CTCF values recorded for hydrogenotrophic probes on day 240 compared with day 125.

The *Crenarchaeota* were also evident in DGGE and proteomic results, with a protein associated with DNA repair (Saci_0192) assigned to a *Sulfolobus* species. Also, a single DGGE band prominent in cDNA profile evident throughout the trial was also assigned to this phylum, suggesting their functional importance in these bioreactors operated at sub-mesophilic conditions.

4.5 Conclusions

This polyphasic approach illustrated the importance of acetoclastic methanogenesis in a well functioning LTAD system, but also showed that at sub-mesophilic conditions, hydrogenotrophic methanogenesis is actively occurring as well. This observation was confirmed through the addition of metaproteomics and RNA profiling in DGGE analysis as well as through the ecophysiological MAR-FISH approach.

The high level of treatment efficiency recorded after the temperature drop (especially in R1) suggests that the mesophilic sludge adapted quickly to low-temperature operation through discrete changes in both bacterial and archaeal community structure. These changes in structure may have occurred primarily within each functional group. For example, although the *Proteobacteria* remained dominant in bacterial communities for both temperatures, clear phylogenetic changes were recorded within this phylum when both samples were compared.

Also, the metabolic flexiblity of *Methanosaeta* was highlighted, with a possible energy conservation system identified as being important in sub-mesophilic growth. Much work is still to be undertaken in LTAD microbial ecology, for example, applying a polyphasic approach to further understand the bacterially mediated process of hydrolysis, a rate limiting step in AD, with the functional analysis of important microbial groups required to further LTAD technology for industrial scale application.

4.6 References

Abram, F., Gunnigle, E. and O'Flaherty, V. (2009) Optimisation of protein extraction and 2-DE for metaproteomics of microbial communities from anaerobic wastewater treatment biofilms. *Electro*. **30**(23), 4149-4151.

Abram, F., Enright, A.M., O'Reilly, J., Botting, C.H., Collins, G., and O'Flaherty, V. 2011. A metaproteomic approach gives functional insights into anaerobic digestion. *J*. *Appl*. *Microbiol*. **110**, 1550-1560.

Andreasen, K. and Nielsen, P. H. (1997) Application of microautoradiography to the study of substrate uptake by filamentous microorganisms in activated sludge. *Appl*. *Environ*. *Microbiol*. **63**, 3662–3668.

APHA. 1998. Standard methods for the examination of water and wastewater In: Clesceri, L.S., Greenberg, A.E., Eaton, A.D (Eds.), Washington, DC, U.S.A.

de Bok, F.A.M., Plugge, C.M. and Stams, A.J.M. (2004) Interspecies electron transfer in methanogenic propionate degrading consortia. *Water Res*. **38**, 1368–1375.

Carrigg, C., Rice, O., Kavanagh, S., Collins, G. and O'Flaherty, V. (2007) DNA extraction method affects microbial community profiles from soils and sediment. *App*. *Microbiol*. *Biotechnol*. **77**(4), 955- 964.

Chistoserdova, L., Vorholt, J., Thauer, R.K. and Lidstrom, M.E. (1998) C1 transfer enzymes and coenzymes linking methylotrophic bacteria and methanogenic archaea. *Sci*. **281**, 99-102.

Coates, J.D., Coughlan, M.F. and Colleran, E. (1996) Simple method for the measurement of the hydrogenotrophic methanogenic activity of anaerobic sludges. *J*. *Microbiol*. *Methods*. **26,** 237-246.

Colleran, E., Concannon, F., Golden, T., Geoghegan, F., Crumlish, B., Killilea, E., Henry, M. and Coates, J. (1992) Use of methanogenic activity tests to characterize anaerobic sludges, screen for anaerobic biodegradability and determine toxicity thresholds against individual anaerobic trophic groups and species. *Water Sci*. *Technol*. **25(7),** 31-40.

Collins, G., McHugh, S., Kearney, A. and O'Flaherty, V. (2004) Psychrophilic anaerobic digestion of a range of wastewaters: process technology, microbial activity and population dynamics. *Water Intell*. *Online*. **40**(6).

Collins, G., Foy, C., McHugh, S., Mahony, T. and O'Flaherty, V. (2005) Anaerobic biological treatment of phenolic wastewater at 15-18°C. *Water Res*. **39**, 1614-1620.

Connaughton, S., Collins, G. and O'Flaherty, V. (2006) Psychrophilic and mesophilic anaerobic digestion of brewery effluent: a comparative study. *Water Res*. **40**, 2503 2510.

Corgié, S.C., Beguiristain, T. and Leyval, C. (2006) Profiling 16S bacterial DNA and RNA: Difference between community structure and transcriptional activity in phenanthrene polluted sand in the vicinity of plant roots. *Soil Biol*. *Biochem*. **38**, 1545-1553.

Cui, Y., Jiang, T., Guo, Z., Zhang, L. and Zhao, L. (2007) Development of anaerobic baffled reactor for domestic wastewater acidification under low-temperature. *Progr*. *Environ*. *Science and Technol*. I, Science Press, Beijing. 810-814.

DeLong, E.F. (1992) Archaea in coastal marine sediments. *Proc*. *Natl*. *Acad*. *Sci*. *U*.*S*.*A*. **89**, 5685-5689.

Dohanyos, M., Zabranska, J., Kutil, J. and Jenicek, P. (2004) Improvement of anaerobic digestion of sludge. *Water Sci*. *Technol*. **49**(10), 89-96.

Enright, A.M., McHugh, S., Collins, G. and O'Flaherty, V. (2005) Low-temperature anaerobic biological treatment of solvent-containing pharmaceutical wastewater. *Water Res*. **39**, 4587-4596.

Enright, A.M., Collins, G. and O'Flaherty, V. (2007) Temporal microbial diversity changes in solvent-degrading anaerobic granular sludge from low-temperature (15°C) wastewater treatment bioreactors. *System*. *App*. *Microbiol*. **30**, 471-482.

Enright, A.M., McGrath, V., Gill, D., Collins, G. and O'Flaherty, V. (2009) Effect of seed sludge and operation conditions on performance and archaeal community structure of low-temperature anaerobic solvent-degrading bioreactors. *Syst*. *Appl*. *Microbiol*. **32**, 65-79.

Goenrich, M., Thauer, R.K., Yurimoto, H. and Kato, N. (2005) Formaldehyde activating enzyme (Fae) and hexulose-6-phosphate synthase (HPS) in *Methanosarcina barkeri*: a possible function in ribose-5-phophate biosynthesis. *Arch*. *Microbiol*. **184**, 41-48.

Gray, N.D., Sherry, A., Grant, R.J., Rowan, A.K., Hubert, C.R.J., Callbeck, C.M., Aitken, C.M., Jones, D.M., Adams, J.J., Larter, S.R. and Head, I.M. (2011) The quantitative significance of *Syntrophaceae* and syntrophic partnerships in methanogenic degradation of crude oil alkanes. *Environ*. *Microbiol*. **13**, 2957-2975.

Griffiths, R.I., Whiteley, A.S., O'Donnell, A.G. and Bailey, M.J. (2000) Rapid method for co-extraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. *Appl*. *Environ*. *Microb*. **66**, 5488-5491.

Hattori, S., Galushko, A.S., Kamagata, Y., and Schink, B. (2005) Operation of the CO dehydrogenase/acetyl coenzyme A pathway in both acetate oxidation and acetate formation by the syntrophically acetate-oxidizing bacterium *Thermacetogenium phaeum*. *J*. *Bacteriol*. **187**, 3471–3476

Hendrickson, E.L. and Leigh, J.A. (2008) Roles of coenzyme F420-reducing hydrogenases and hydrogen- and F_{420}-dependent methylenetetrahydromethanopterin dehydrogenases in reduction of F_{420} and production of hydrogen during methanogenesis. *J*. *Bacteriol*. **190**, 4818–4821.

Kato, M., Field, J.A., Versteeg, P. and Lettinga, G. (1994) Feasibility of the expanded granular sludge bed (EGSB) reactors for the anaerobic treatment of low strength soluble wastewaters. *Biotechnol*. *Bioengineer*. **44**, 469-479.

Kato, M.T., Field, J.A. and Lettinga, G. (1997) The anaerobic treatment of low strength wastewaters in UASB and EGSB reactors. *Water Sci*. *Tech*. **36**(6), 375-382.

Kotsyurbenko, O. R. (2005) Trophic interactions in the methanogenic microbial community of low-temperature terrestrial ecosystems. *FEMS Microb*. *Ecol*. **53**(1), 3-13.

Lane, D.J., Pace, B., Olsen, G.J., Stahl, D.A., Sogin, M.L. and Pace, N.R. (1985) Rapid determination of 16S ribosomal RNA sequences for phylogenetic analysis. *Proc*. *Natl*. *Acad*. *Sci*. *U*.*S*.*A*. **82**, 6955-6959.

Lee, C., Kim, J., Hwang, K., O'Flaherty, V. and Hwang, S. (2009) Quantitative analysis of methanogenic community dynamics in three anaerobic batch digesters treating different wastewaters. *Water Res*. **43**, 157-165.

Lettinga, G., Rebac, S., Parshina, S., Nozhevnikova, A., van Lier, J.B. and Stams, A.J.M. (1999) High-rate anaerobic treatment of wastewater at low-temperatures. *Appl*. *Environ*. *Microbiol*. **65**, 1696-1702.

Martinez-Sosa, D., Helmreich, B., Netter, T., Paris, S., Bischof, F. and Harald, H. (2011) Anaerobic submerged membrane bioreactor (AnSMBR) for municipal wastewater treatment under mesophilic and psychrophilic temperature conditions. *Biores*. *Technol*. **22**, 10377-10385.

McInerney, M.J., Voordouw, G., Jenneman, G.E. and Sublette, K.L. (2007) Oil field microbiology. In ''Manual of Environmental Microbiology'' (C. J. Hurst, R. L. Crawford, J.L., Garland, D.A., Lipson, A.L., and Stetzenbach, L.D. Eds.), 898–911. ASM, Washington, DC.

McKeown, R., Scully, C., Enright, A.M., Chinalia, F.A., Lee, C., Mahony, T., Collins, G., O'Flaherty, V., 2009. Psychrophilic methanogenic community development during long-term cultivation of anaerobic granular biofilms. *ISME*. *J*. **3**, 1231-1242.

McKeown, R., Hughes, D., Collins, G., Mahony, T. and O'Flaherty, V. (2012) Low-temperature anaerobic digestion for wastewater treatment. *Biotechnol*. **23**, 444-451.

Nielsen, P.H. and Nielsen, J.L. (2005) Microautoradiography: recent advances within the studies of the ecophysiology of bacteria in biofilms. *Water Sci*. *Technol*. **53**(7), 187-194.

Okabe, S., Kindaichi, T. and Ito, T. (2005) Fate of ^{14}C-labelled microbial products derived from nitrifying bacteria in autotrophic nitrifying biofilms. *Appl*. *Environ*. *Microbiol*. **71**, 3987–3994.

O'Reilly, J., Lee, C., Collins, G., Chinalia, F.A., Mahony. T. and O'Flaherty, V. (2009) Quantitative and qualitative analysis of methanogenic communities in mesophilically and psychrophilically cultivated anaerobic granular biofilms. *Water Res*. **43**, 3365-3374.

Ramirez, I., Volcke, E.I.P., Rajinikanth, R. and Steyer, J.P. (2009) Modeling microbial diversity in anaerobic digestion through an extended ADM1 model. *Water Res*. **43**, 2787–2800

Schocke, L. and Schink, B. (1997) Energetics of methanogenic benzoate degradation by Syntrophus gentianae in syntrophic coculture. *Microbiol*. **143**, 2345 - 2351

Schramm, A., de Beer, D., Wagner, M. and Amann, R. (1998) Identification and activity in situ of *Nitrosospira* and *Nitrospira* spp. as dominant populations in a nitrifying fluidized bed reactor. *Appl*. *Environ*. *Microbiol*. **64**, 3480-3485.

Sekiguchi, Y., Kamagata, Y., Nakamura, K., Ohashi, A. and Harada, H. (1999) Fluorescence in situ hybridization using 16S rRNA- targeted oligonucleotides reveals localization of methanogens and selected uncultured bacteria in mesophilic and thermophilic sludge granules. *Appl*. *Environ*. *Microbiol*. **65**, 1280–1288.

Shelton, D.R. and Tiedje, J.M. (1984) General method for determining anaerobic biodegradation potential. *Appl*. *Environ*. *Microbiol*. **47,** 850-857.

Shevchenko, A., Wilm, M., Vorm, O. and Mann, M. (1996) Mass spectrometric sequencing of proteins from silver- stained polyacrylamide gels. *Anal*. *Chem*. **68**, 850–858.

Shilov, I.V., Seymour, S.L., Patel, A.A., Loboda, A., Tang, W.H., Keating, S.P., Hunter, H.L., Nuwaysir ,L.M. and Schaeffer D.A. (2007) The Paragon algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol. Cell. Proteom.* **6**, 1638–1655.

Siggins, A., Enright, A.M. and O'Flaherty, V. (2011) Low-temperature (7°C) anaerobic treatment of a trichloroethylene-contaminated wastewater: Microbial community development. *Water Res.* **45**, 4035, 4046.

Smith, K.S. and Ingram-Smith, C. (2007). *Methanosaeta*, the forgotten methanogen? *Trends Microbiol.* **15**(4), 150-155.

Soderberg, T. (2005) Biosynthesis of ribose-5-phosphate and erythrose-4-phosphate in archaea: a phylogenetic analysis of archaeal genomes. *Arch.* **1**, 347-352.

Stackebrandt, E. and Goodfellow M. (1991) Nucleic Acid Techniques in Bacteria Systematics. Chichester: Wiley.

Stamatakis, A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinform.* **22**(21), 2688–2690.

Syutsubo, K., Yoochatchaval, W., Yoshida, H., Nishiyama, K., Okawara, M., Sumino, H., Araki, N., Harada, H. and Ohashi, A. (2008) Changes of microbial characteristics of retained sludge during low-temperature operation of an EGSB reactor for low strength wastewater treatment. *Water Sci. Tech.* **57**(2), 277-281.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**(24), 4876-4882.

Viraraghavan,,T. and Kikkeri, S.R. (1990). Effect of temperature on anaerobic filter treatment of dairy wastewater. *Water Sci. Technol.* **22**, 191–198.

Wallner, G., Amann, R. and Beisker, W. (1993) Optimising fluorescent in situ hybridization of suspended cells with rRNA-targeted oligonucleotide probes for the flow cytometric identification of microorganisms. *Cytom.* **14**, 136–143.

Welte, C. and Deppenmeier, U. (2011) Membrane-bound electron transport in *Methanosaeta thermophila*. *J. Bacteriol.* **193**, 2868–2870.

World Health Organisation, (2008) Guidelines for drinking-water quality. ISBN 978 92 4 154761 1

Verstraete, W. and Vandevivere. (1999) New and broader applications of anaerobic digestion. *Crit. Rev. Environ. Sci. Technol.* **29**(2), 151-173.

Yu, Y., Lee, C., Kim, J. and Hwang, S. (2005) Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnol*. *Bioeng*. **89**, 670-679.

Chapter 5

Final Conclusions and Recommendations

5.1 Conclusions

The results of this thesis demonstrated the value of employing a functional-based polyphasic approach to characterise the mixed microbial communities present in LTAD bioreactors. Indeed, the results gathered implied that the sub-mesophilic anaerobic treatment of a VFA-based wastewater was facilitated by the psychrotolerant capabilities of specific organisms with optima in the mesophilic temperature range, while low-temperature mediated successional development of particular microbial groups was also evident.

The acetoclastic *Methanosaeta* were the most prominent methanogenic archaeal group identified in both LTAD trials (Chapter 2 & Chapter 4). In end-point biomass samples (Chapter 2), this group comprised 76% (15°C) and 91% (7°C) of archaeal clones, with low-temperature biomass after a temperature drop from mesophilic conditions (Chapter 4) having 45% of clone library closely aligned with *Methanosaeta*-like organisms. This dominance was reflected in qPCR analysis. This is in agreement with previous studies, where dominance of *Methanosaeta* sp. was reported in stable bioreactors with low acetate concentrations, noted during the majority of bioreactor study through the analysis of effluent VFA levels (Chapter 4). This level of detection in clone library and qPCR analysis was in agreement with proteomic analysis, with the highest amount of assigned proteins relating to this group (69% - Chapter 2; 66% Chapter 4). What was intriguing about the proteomic coverage of *Methanosaeta* sp. functioning in low-temperature AD biomass, was the possible uncovering of particular adaptation strategies this methanogenic group employs, which have not been characterised in an AD context to our knowledge. This included the finding of two acetyl CoA synthetase proteins relating to separate gene loci, with a particular protein found to be upregulated in both 15°C and 7°C biomass in comparison to mesophilic biomass (Chapter 2), while the upregulation of a F_{420} hydrogenase protein usually affiliated with methylotrophic methanogenesis, may give a direct example of sub-mesophilic energy conservation for this methanogenic group (Chapter 4).

Although the dominance of *Methanosaeta* was clear in biomass underpinning both bioreactor trials, the use of a PCR-based and PCR-independent approaches allowed the elucidation of sensative communty succession profiles relating to other methanogenic groups functioning in LTAD also. As with previous studies, the emergence of hydrogenotrophic methanogenesis in conjunction with decreased operation temperature was recorded, with higher 15°C SMA values from biomass sampled at the end of trial compared with biomass sampled before the temperature drop (Chapter 4). Previous studies have recorded an increase *Methanocorpusculum* sp. with the decrease in operating temperature through clone library and qPCR analysis, suggesting that an increase in abundance correlated with functional importance. Although an increase in phylogenetic representation was recorded in an end-point biomass sample at 15°C (19% archaeal clones; Chapter 2), the proteomic profile suggested that *Methanobacteriales* sp. were more important for methanogenesis, highlighting the importance of incorporating a functional-based approach in LTAD.

Furthermore, the polyphasic approach employed in chapter 4 demonstrated that genetic fingerprinting methodologies such as DGGE can provide important information relating to community dynamics, especially when there is a variable such as temperature. This was demonstrated through the incorporation of RNA-based analysis in conjunction with a general DNA-based DGGE approach. If DNA-based DGGE profiles along with clone library and qPCR analysis were the only methodologies applied in this study, key functional groups may have been overlooked. Indeed, metaproteomic and RNA-based DGGE profiling facilitated an more sensative approach in uncovering the key functional groups underpinning these LTAD granular biofilms. For example, *Methanospirillum*, belonging to the order *Methanomicrobiales*, was not detected in archaeal clone libraries, however, several proteins were identified relating to this hydrogenotrophic methanogen, with a protein relating to methanogenesis upregulated in low-temperature biomass. Also, two DGGE sequences were assigned to this group, relating to bands in RNA-based profiles more prominent at low-temperature. In chapter 2, a *Methanospirillum* protein relating to methanogenesis was upregulated in low-temperature biomass samples compared with mesophilic sample, highlighting through replication, the

successful application of methodologies employed in this research.

This approach also highlighted the possible community succession within a particular group. For example, in both bioreactor studies the methanogenic order *Methanobacteriales* were primarily represented in the metaproteomic profiles through numerous species assigned to the same protein, methyl coenzyme M reductase, responsible for the final step in methanogenic formation of CH_4. However, the level of expression differed between particular species in relation to sample. For example, this protein belonging to the generally classified thermophile *Methanothermobacter thermautotrophicus* was found to be upregulated in the end-point 15°C sample (Chapter 2) and at the end of the second bioreactor trial (Chapter 4) when compared to the corresponding mesophilic samples, while the reverse was recorded for *M. marburgensis* (Chapter 2) and an unclassified *Methanobacteriales* sp. (Chapter 4). It is therefore apparent that successive changes within functional groups (such as *Methanobacteriales*) in relation to temperature may facilitate LTAD treatment without significant impact on operational performance, as observed in Chapter 4.

Furthermore, in order to go further into uncovering low-temperature methanogenic adaptation strategies, we applied a proteomic approach on a *Methanosarcina* sp. (Chapter 3) after a protein assigned to this group was identified at a significant level in low-temperature biomass (Chapter 2). Interestingly, this study presented the possibility that this methanogen could proliferate in sub-mesophilic conditions after a period of 'adaptation', where low cell numbers were in coordination with high cell viability, giving important insights into how environmental homologs survive and function in LTAD communities. For example, through the regulation of elongation factor proteins and chaperone protein expression during adptation to low-temperature conditions. Although the second bioreactor trial (Chapter 4) did not return any proteins assigned to this methanogenic group, a DGGE sequence was identified relating to this group, associated with a band detected only in RNA-based profiles from samples taken directly after the temperature drop to 15°C. This, in coordination with an increase in acetate concentration recorded through VFA analysis in both bioreactors during these sampling points, suggests that

Methanosarcina were able to proliferate for a short period of time before acetate levels dropped to pre-temperature drop concentrations. Once again, the comparative results gathered from two separate approaches (VFA concentration, RNA-based DGGE profiling) confirm the importance of applying a polyphasic approach to LTAD when coordinating microbial community dynamics.

Although the analysis of proteins from mixed communities in LTAD is still relatively novel, with improved methodologies for protein extraction and identification still in development, this study represents a concise body of work which shows that a functional-based polyphasic approach does work and can allow sensative community information to be gathered, which may be missed by employing traditional community profiling techniques.

5.2 Future recommendations

Although the results presented in this thesis are concise and have given insights into the grey area of microbial community functioning in LTAD systems, significant knowledge gaps in the field of LTAD remain. Therefore, the following recommendations are proposed:

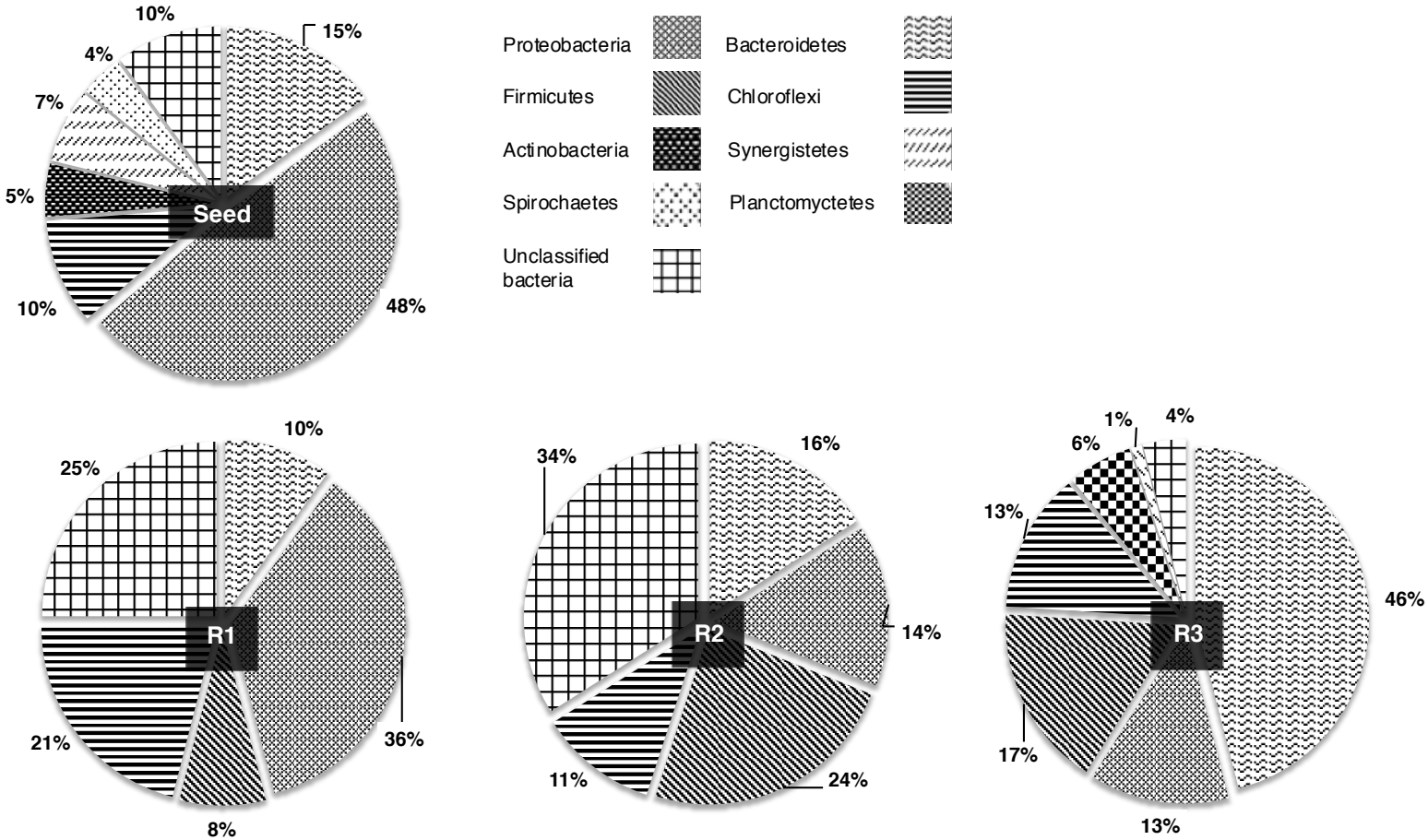
- The use of RNA-based DGGE profiling in coordination with DNA-based methods for the community profiling of future LTAD studies, particularly in temperatures below 10C, where psychrophilic communities may become more prevalent, and as such important trophic information may be recorded.
- Metaproteomic application for screening seed biomass prior to LTAD application, with potential for targeting important functional groups and particular characteristics.
- RNA– based qPCR quantification of methanogenic methyl-coenzyme M gene, which could be employed as an improved accuracy measurement for methanogenic groups, directly relating to function.
- Pure culture characterisation of the functionally important methanogenic groups, such as *Methanosaeta*, at low-temperatures. Although well documented that *Methanosaeta* are difficult to grow and maintain in pure culture, important insights into the metabolic capacity of this organism in LTAD is essential from a biotechnological point of view.

Appendix 1: Characteristics of primers and probes used in this study. F - Forward primer; R - Reverse primer; P - Probe.

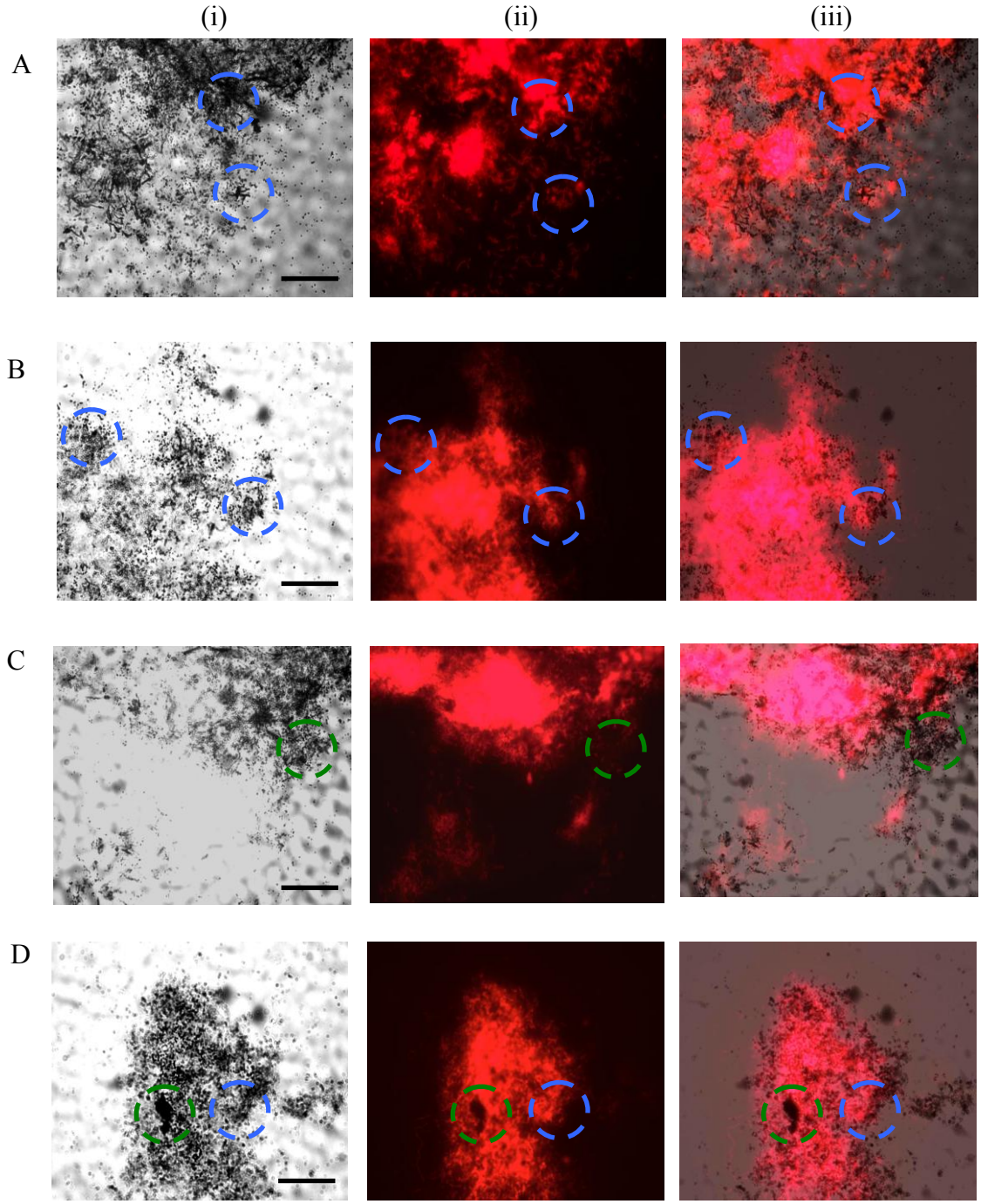
Molecular analysis	Target group	Primer/Probe Sequence (5' → 3')		Amplicon size (bp)	Reference	Relevant chapter
Clone libraries	Bacteria	27F	AGA GTT TGA TCC TGG CTC AG	1433	DeLong, (1992)	
		1392R	ACG GGC GGT GTG TRC		Lane *et al.* (1985)	
	Archaea	21F	TTC CGG TTG ATC CYG CCG GA	920	Stackebrandt & Goodfellow, (1991)	
		958R	YCC GGC GTT GAM TCC AAT T		DeLong, (1992)	
DGGE	Archaea	787F	ATTAG ATACC CSBGT AGTCC	273	Yu *et al*. (2005)	
		1059R	GCCAT GCACC WCCTC T		Yu *et al*. (2005)	
qPCR	Methanobacteriales (MBT set)	857F	CGW AGG GAA GCT GTT AAG T	343	Yu *et al*. (2005)	
		1059R	TAC CGT CGT CCA CTC CTT		Yu *et al*. (2005)	
		929P	AGC ACC ACA ACG CGT GGA		Yu *et al*. (2005)	
	Methanomicrobiales (MMB set)	282F	ATC GRT ACG GGT TGT GGG	506	Yu *et al*. (2005)	
		832R	CAC CTA ACG CRC ATH GTT TAC		Yu *et al*. (2005)	
		749P	TYC GAC AGT GAG GRA CGA AAG CTG		Yu *et al*. (2005)	

Molecular analysis	Target group	Primer/Probe Sequence (5' → 3')		Amplicon size (bp)	Reference	Relevant chapter
qPCR	Methanosaetaceae (MSt set)	702F	GAA ACC GYG ATA AGG GGA	164	Yu *et al*. (2005)	
		862R	TAG CGA RCA TCG TTT ACG		Yu *et al*. (2005)	
		753P	TTA GCA AGG GCC GGG CAA		Yu *et al*. (2005)	
	Methanosarcinaceae (MSc set)	380F	TAA TCC TYG ARG GAC CAC CA	408	Yu *et al*. (2005)	
		828R	CCT ACG GCA CCR ACM AC		Yu *et al*. (2005)	
		492P	ACG GCA AGG GAC GAA AGC TAG G		Yu *et al*. (2005)	

Appendix 2: Comparative distribution of bacterial groups detected by clone library analysis performed on seed biomass and bioreactor biomass on days 631 for R1 (37°C) and R2 (15°C) and day 609 for R3 (7°C). Chapter 2.



Appendix 2: Selected MAR-FISH images after 15 hours of incubation with [^{14}C] acetic acid on (A) day 125 and (B) day 240 and H_2, sodium-[^{14}C]-bicarbonate on (C) day 125 and (D) day 240. For each sample, (i) MAR (silver grains), (ii) FISH signal and (iii) MAR-FISH overlay are shown, depicting MAR-positive cells (indicated by dashed blue line) and other areas of substrates uptake (dashed green line). Scale bar is 10 µm.



Appendix 4: Relevant information for partial archaeal 16S rRNA gene sequences retrieved from DGGE bands analysed in Chapter 4.

Accession number	Affiliated band present in sample profile										Closeste relative (BLASTn)
	Day 125 DNA	Day 125 cDNA	Day 126 DNA	Day 126 cDNA	Day 128 DNA	Day 128 cDNA	Day 132 DNA	Day 132 cDNA	Day240 DNA	Day 240 cDNA	
KC305601	O	O		O		O		O	O	O	*Methanosaeta concilii* (NR028242)
KC305602	O		O	O	O	O	O	O	O	O	*Methanosaeta concilii* (NR028242)
KC305603	O		O	O	O	O	O	O	O	O	*Methanobacterium beijingense* (NR028202)
KC305604	O	O	O	O	O	O	O	O	O	O	*Methanosaeta concilii* (NR028242)
KC305605	O		O	O	O		O	O	O		*Methanobacterium formicicum* (NR025028)
KC305606	O	O	O	O	O	O	O	O	O	O	*Methanosaeta concilii* (NR028242)
KC305607		O				O				O	*Methanobacterium beijingense* (NR028202)
KC305608		O		O		O		O		O	*Thermofilum pendens* (NR029214)

Accession number	Affiliated band present in sample profile										Closeste relative (BLASTn)
	Day 125 DNA	Day 125 cDNA	Day 126 DNA	Day 126 cDNA	Day 128 DNA	Day 128 cDNA	Day 132 DNA	Day 132 cDNA	Day240 DNA	Day 240 cDNA	
KC305609	O	O		O		O		O		O	*Methanospirillum hungatei* (NR042789)
KC305610	O	O	O	O	O	O	O	O	O	O	*Methanolinea tarda* (NR028163)
KC305611	O		O	O			O				*Methanofollis liminatans* (NR028254)
KC305612	O	O	O	O	O	O	O			O	*Methanobacterium beijingense* (NR028202)
KC305613	O		O		O		O		O		*Methanolinea tarda* (NR028163)
KC305614		O		O		O		O		O	*Ignisphaera aggregans* (NR043512)
KC305615		O		O		O		O	O	O	*Methanobacterium beijingense* (NR028202)
KC305616				O		O		O			*Methanosarcina mazei* (NR041956)

Accession number	Affiliated band present in sample profile										Closeste relative (BLASTn)
	Day 125 DNA	Day 125 cDNA	Day 126 DNA	Day 126 cDNA	Day 128 DNA	Day 128 cDNA	Day 132 DNA	Day 132 cDNA	Day240 DNA	Day 240 cDNA	
KC305617		O		O		O		O		O	*Methanolinea tarda* (NR028163)
KC305618		O		O		O		O		O	*Methanolinea tarda* (NR028163)
KC305619		O		O		O		O			*Methanospirillum hungatei* (NR042789)
KC305620		O				O					*Methanobacterium formicicum* (NR025028)
KC305621	O	O	O	O	O	O	O	O	O	O	*Methanothermobacter thermautotrophicus* (NR028241)
KC305622	O		O		O		O		O	O	*Methanobacterium beijingense* (NR028202)
KC305623	O	O	O	O	O	O	O	O	O	O	*Methanosaeta concilii* (NR028242)

Electrophoresis 2009, *30*, 4149–4151

Florence Abram[1]
Eoin Gunnigle[1]
Vincent O'Flaherty[1,2]

[1]Department of Microbiology, School of Natural Sciences, National University of Ireland, Galway, Ireland
[2]Energy Research Centre, Environmental Change Institute, National University of Ireland, Galway, Ireland

Received July 24, 2009
Revised September 10, 2009
Accepted September 10, 2009

Short Communication

Optimisation of protein extraction and 2-DE for metaproteomics of microbial communities from anaerobic wastewater treatment biofilms

The feasibility of metaproteomic analysis of the microbial communities underpinning anaerobic wastewater treatment relies on efficient protein extraction and separation techniques. The microorganisms involved in anaerobic digestion (AD) of wastewater typically aggregate to form tightly organised, spherical granules, from which protein extraction is challenging. Here, we compare two methods of protein extraction, one using a French press previously used successfully to analyse the proteome of an activated sludge [Wilmes, P., Bond, P. L., *Environ. Microbiol.* 2004, *6*, 911–920.] and one using sonication developed in the context of pure culture [Abram, F., Wan-Lin, S., Wiedmann, M., Boor, K. J., Coote, P., Botting, C., Karatzas, K. A. G., O'Byrne, C. P., *Appl. Environ. Microbiol.* 2008, *74*, 594–604.]. We used 2-DE to carry out this comparison. The protein extraction using the sonication method resulted in a significant increase in the number of protein spots and higher quality 2-D gels.

Keywords:
Anaerobic digestion / Metaproteomics / Sludge granules / Wastewater treatment.
DOI 10.1002/elps.200900474

Many years after its introduction by O'Farrell in 1975 [1], 2-DE continues to play an important role in the experimental analysis of proteins and protein mixtures. Since the first publication of genome sequences in 1995 [2], the most commonly applied approach for proteomic analysis is performed by combining 2-DE, to separate and visualise proteins, with subsequent protein identification by MS [3]. Where mixed microbial communities prevail, the major challenge of the post-genomic era is to define the link between microbial community composition and function. The term metaproteomics was proposed in 2004 [4] as the "large-scale characterisation of the entire protein complement of environmental microbiota at a given point in time". An efficient protein extraction method is inherent to the subsequent analysis of proteins by 2-DE. Extracting proteins from anaerobic mixed consortia used in wastewater treatment is proving to be particularly complex.

Anaerobic digestion (AD) has become an essential and well-established technology to treat various problematic wastewaters. The natural process of AD is driven by complex interactions between several trophic groups of different microorganisms, which sequentially and co-operatively convert complex organic molecules to biogas [5]. The microorganisms involved in engineered AD systems for waste treatment and bioenergy production aggregate to form spherical biofilms (anaerobic sludge granules; diameter, *c.* 0.5–2 mm). Individual granules each theoretically contain all of the trophic groups required for the substrate depletion in question. Although AD is widely applied, relatively little is known about the behaviour of the mixed populations in anaerobic sludge and the further development and improved exploitation of this biotechnology is contingent on our better understanding of the microbial ecology of those consortia. We propose to apply, for the first time, a metaproteomic approach to investigate protein expression in mixed communities underpinning anaerobic bioreactors. To this end, we developed an efficient protein extraction method to ensure a successful protein separation by 2-DE. We compared two extraction procedures, one using a French press adapted from Wilmes and Bond [4] and one using sonication adapted from Abram *et al.* [6]. We collected 50 mL anaerobic sludge granules, from which proteins were extracted by following the protocols outlined below.

French press protocol (adapted from Wilmes and Bond [4]*)*: Immediately after sampling, the sludge granules were immersed in liquid nitrogen. Using a mortar and pestle, the granules were ground to a fine powder under freezing conditions. The pulverised samples were resuspended in 50 mL of 0.9% w/v NaCl and centrifuged at 4500 × *g* for 20 min at 4°C. The resulting pellets were washed twice in 50 mL 50 mM Tris-HCl (pH 7.0) by centrifugation at

Correspondence: Professor Vincent O'Flaherty, Microbial Ecology Laboratory, Department of Microbiology, School of Natural Sciences, National University of Ireland, Galway, University Road, Galway, Ireland
E-mail: vincent.oflaherty@nuigalway.ie
Fax: +353-91-49-4598

Abbreviation: AD, anaerobic digestion

4500 × *g* for 10 min at 4°C. The pellets were resuspended in 10 mL extraction buffer (7 M urea, 2 M thiourea, 4% v/v Triton X-100, 10 mM Tris-base, 1 mM EDTA, 50 mM DTT and 10 μL protease inhibitor cocktail (Amersham) *per* mL). Suspensions were left on ice for 2 h, with vortexing every 15 min, and were subsequently passed three times through a French press (AB BIOX, Stockholm, Sweden) at 28 000 psi. Soluble proteins were separated from the cell debris by centrifugation for 30 min at 13 000 × *g*. Benzonase (44 U/mL) was added to each sample, which were then incubated at 37°C for 30 min. The benzonase treatment was repeated once and the protein samples were centrifuged at 13 000 × *g* for 30 min. Supernatants were stored at −80°C before analysis by 2-DE.

Sonication protocol (adapted from Abram et al. [6]*)*: Immediately upon sampling, granules were washed twice by centrifugation at 8000 × *g* for 15 min at 4°C in sonication buffer (10 mM Tris-base, 0.1 mM EDTA and 5 mM $MgCl_2$). The pellets were then frozen in liquid nitrogen and lyophilised. The pulverised samples were resuspended in 10 mL sonication buffer to which 10 μL protease inhibitor cocktail *per* mL and 10 μL 10% w/v SDS *per* mL were added. Suspensions were sonicated on ice using an MSE soniprep 150 at an amplitude of 22 μm for 30 s. Fifteen pulses were applied, with 30 s rest intervals between pulses. Soluble proteins were separated from cell debris by centrifugation and benzonase treatments were carried out as described above in the French press protocol. The protein supernatants were stored at −80°C.

Protein concentrations were determined using a BioRad RC DC kit with bovine serum albumin as a standard following the manufacturers' protocol. For each procedure protein standards were prepared in the same buffer as protein samples and adequate controls and blanks (extraction buffer or sonication buffer only) were assayed. In addition, for the French press protocol a second wash step was included in the procedure as recommended by the manufacturer. The French press procedure resulted in higher protein concentrations than the sonication method with typical yields of about 11 and 3 mg/mL, respectively. Eight hundred micrograms of each protein extract were precipitated using a 2-D Clean-Up kit (Amersham Biosciences). The dried protein pellets were resuspended in 135 μL IEF rehydration solution (Amersham Biosciences) to which IPG buffer (final concentration, 0.2% v/v; Amersham Biosciences) and 2.5 mg DTT *per* mL had been freshly added. The samples were incubated at room temperature for 2 h with vigorous vortexing every 10 min. The suspensions were then centrifuged at 18 000 × *g* for 10 min to remove any insoluble material. IPG strips (pH 4–7; linear; 7 cm; Amersham Biosciences) were rehydrated overnight with the rehydration solution containing the protein samples to be analysed. Gels were run with a Protean IEF cell unit (BioRad) at 250 V for 20 min, followed by 4000 V for 2 h and 4000 V for 10 000 V/h. Gels strips were subsequently equilibrated for 20 min in equilibration buffer A (6 M urea, 30% v/v glycerol, 2% w/v SDS, 20 mg DTT *per* mL 0.05% w/v bromophenol blue, 45 mM Tris-base [pH 7]) followed by 20 min in equilibration buffer B (as buffer A but containing 25 mg iodoacetamide *per* mL instead of DTT). The second-dimension gels were performed using 12% v/v acrylamide and were run in pairs at 70 mA using a Mini-Protean cell electrophoresis unit (BioRad). Molecular weight markers ranging from 10 to 225 kDa (Promega) were run in the second dimension as size standards. Gels were stained overnight with GelCode Blue Safe Protein Stain (Thermo Scientific) and then destained in deionised, distilled water for several hours. Gel images were captured by scanning the gels with an Epson Perfection V350 photo scanner at a resolution of 800 dpi. For each extraction procedure, eight gels were run corresponding to two duplicate independent extractions and four technical replicates. Gels were analysed with PDQuest-Advanced software, version 8.0.1 (BioRad). Spot counts were obtained using the spot detection wizard enabling the Gaussian model option as recommended by the manufacturer.

The sonication protocol improved considerably the quality of the 2-D gels obtained when compared to those generated by the French press protocol, as illustrated in Fig. 1. The sonication extraction procedure produced gels displaying an average of 368 spots (SD 20.4) against an average of 21 spots (SD 2.2) when protein were extracted using the French press protocol. Although the French press protocol led to higher protein concentrations than the sonication procedure, it is possible that interfering substances co-extracted during the French press protocol could artificially increase the value obtained by the protein concentration determination assay. It was previously reported that after multiple passes through a high pressure homogeniser, the cell debris might be degraded to smaller fragments resulting in less effective downstream separation [7]. If the higher protein concentrations obtained by the French press protocol was artifactual, the low number of protein spots present on the 2-D gels could be a result of an insufficient protein load. To test this, IEF strips were run with protein loads of up to 4 mg, which did not result in any increase in the number of spots on the gels (data not shown). We also tested the French press protocol using the sonication buffer (10 mM Tris-base, 0.1 mM EDTA and 5 mM $MgCl_2$, 10 μL protease inhibitor cocktail (Amersham) *per* mL and 10 μL 10% w/v SDS *per* mL), as well as the sonication protocol using the extraction buffer (7 M urea, 2 M thiourea, 4% v/v Triton X-100, 10 mM Tris-base, 1 mM EDTA, 50 mM DTT and 10 μL protease inhibitor cocktail (Amersham) *per* mL). The change of buffer in the French protocol led to a marginally higher number of spots on the gels with an average of 35 spots per gels (SD 2.6), and the appearance of horizontal streaking (Fig. 2A), typically associated with the presence of non-proteinic contaminants in the protein samples [8, 9]. The use of the extraction buffer in the sonication protocol decreased considerably the quality of the resulting 2-D gels (Fig. 2B; average of 17 spots with a SD of 2.1), which were very similar to those obtained with the French press protocol initially tested [4] (Fig. 1A). This result would indicate that the proteins present in our anaerobic sludge sample display a very limited solubility in the extraction buffer used in the French press protocol originally designed

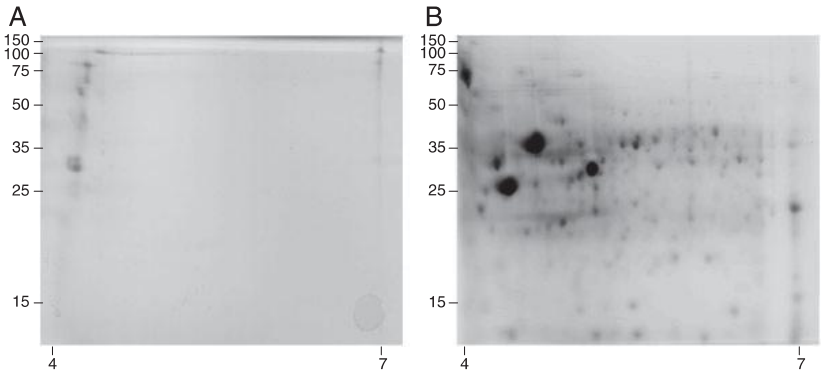


Figure 1. 2-D gels obtained using the French press extraction protocol (A) and the proposed sonication protocol (B). The molecular weight marker values indicated on the left side of the gel images are in kDa and the 7 cm IPG strips used displayed a pH gradient ranged from 4 to 7 (indicated below the gel pictures).

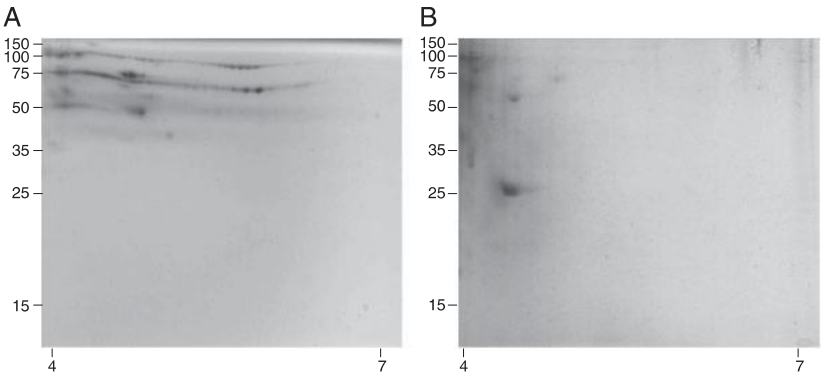


Figure 2. 2-D gels obtained using the French press extraction protocol with the sonication buffer (A) and the proposed sonication protocol with the French press extraction buffer (B). The molecular weight marker values indicated on the left side of the gel images are in kDa and the 7 cm IPG strips used displayed a pH gradient ranged from 4 to 7 (indicated below the gel pictures).

for activated sludge [4]. The flocculated activated sludge investigated by Wilmes and Bond [4] contained mainly bacteria (88.3%) while the granular anaerobic sludge used in the present study is typically dominated by archaeal species (75%; [10]). The sonication method initially tested in this study [6] was successfully applied on various anaerobic sludge samples.

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (Principal Investigator Award and a Research Frontiers Programme Award). We are very grateful to Dr. Gavin Collins for providing critical comments on the manuscript.

The authors have declared no conflict of interest.

References

[1] O'Farrell, P. H., *J. Biol. Chem.* 1975, *250*, 4007–4021.

[2] Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J. *et al.*, *Science* 1995, *269*, 496–512.

[3] Smith, R., *Methods Mol. Biol.* 2009, *519*, 1–16.

[4] Wilmes, P., Bond, P. L., *Environ. Microbiol.* 2004, *6*, 911–920.

[5] McCarty, P. L., *Wat. Sci. Technol.* 2001, *44*, 149–156.

[6] Abram, F., Wan-Lin, S., Wiedmann, M., Boor, K. J., Coote, P., Botting, C., Karatzas, K. A. G., O'Byrne, C. P., *Appl. Environ. Microbiol.* 2008, *74*, 594–604.

[7] Baldwin, C., Robinson, C. W., *Biotechnol. Techn.* 1990, *4*, 329–334.

[8] Bodzon-Kulakowska, A., Bierczynska-Krzysik, A., Dylag, Y., Drabik, A., Suder, P., Noga, M., Jarzebinska, J., Silberring, J., *J. Chromatogr. B* 2007, *849*, 1–31.

[9] Gorg, A., Obermaier, C., Boguth, G., Harder, A., Scheibe, B., Wildgruber, R., Weiss, W., *Electrophoresis* 2000, *21*, 1037–1353.

[10] Enright, A. M., McGrath, V., Gill, D., Collins, G., O'Flaherty, V., *Syst. Appl. Microbiol.* 2009, *32*, 65–79.

FEMS MICROBIOLOGY ECOLOGY

MINIREVIEW

Exploring mixed microbial community functioning: recent advances in metaproteomics

Alma Siggins[1], Eoin Gunnigle[1] & Florence Abram[2]

[1]Microbial Ecology Laboratory, Department of Microbiology and Ryan Institute, National University of Ireland, Galway (NUI, Galway), Galway, Ireland; and [2]Functional Environmental Microbiology, Department of Microbiology, National University of Ireland, Galway (NUI, Galway), Galway, Ireland

Correspondence: Florence Abram, Functional Environmental Microbiology, Department of Microbiology, National University of Ireland, Galway (NUI, Galway), University Road, Galway, Ireland. Tel.: +353 0 91 492390; fax: +353 0 91 494598; e-mail: florence.abram@nuigalway.ie

Received 16 May 2011; revised 7 October 2011; accepted 13 December 2011.

DOI: 10.1111/j.1574-6941.2011.01284.x

Editor: Ian Head

Keywords
environmental proteomics; human gut microbiota; marine and freshwater environment; soil; bioengineered systems.

Abstract

System approaches to elucidate ecosystem functioning constitute an emerging area of research within microbial ecology. Such approaches aim at investigating all levels of biological information (DNA, RNA, proteins and metabolites) to capture the functional interactions occurring in a given ecosystem and track down characteristics that could not be accessed by the study of isolated components. In this context, the study of the proteins collectively expressed by all the microorganisms present within an ecosystem (metaproteomics) is not only crucial but can also provide insights into microbial functionality. Overall, the success of metaproteomics is closely linked to metagenomics, and with the exponential increase in the availability of metagenome sequences, this field of research is starting to experience generation of an overwhelming amount of data, which requires systematic analysis. Metaproteomics has been employed in very diverse environments, and this review discusses the recent advances achieved in the context of human biology, soil, marine and freshwater environments as well as natural and bioengineered systems.

Introduction

Microorganisms occupy virtually every habitat on our planet, and their activities largely determine the environmental conditions of today's world. Indeed, microorganisms are heavily involved in biogeochemistry, ensuring the recycling of elements such as carbon and nitrogen (Madsen, 2011). In addition, microorganisms are extensively used to degrade anthropogenic waste prior to release into the environment (Hussain *et al.*, 2010; Park *et al.*, 2011). In their natural habitat, microorganisms coexist in mixed communities, the complexity of which is specific to each environment, for example from six estimated individual taxa for an acid mine drainage biofilm (Ram *et al.*, 2005), up to 10^6 estimated taxa per gram of soil (Wilmes & Bond, 2006). As most of the microorganisms present in the environment have not been cultured, their investigation requires the use of molecular techniques that bypass the traditional isolation and cultivation of individual species (Amann *et al.*, 1995). Moreover, even when isolation is possible, a single species removed from its natural environment might not necessarily display the same characteristics under laboratory conditions as it does within its ecological niche. Therefore, the study of mixed microbial communities within their natural environment is key to the investigation of the diverse roles played by microorganisms, and to the identification of the microbial potential for biotechnological application, including but not limited to: pharmaceutical, diagnostics, waste treatment, bioremediation and renewable energy generation. An emerging field of research in microbial ecology encompasses system approaches (Fig. 1), whereby all levels of biological information are investigated (DNA, RNA, proteins and metabolites) to capture the functional interactions occurring in a given ecosystem and identify characteristics that could not be accessed by the study of isolated components (Raes & Bork, 2008; Röling *et al.*, 2010). Recent technological advances, including the development of high-throughput 'omics' methods, make such system approaches possible, where mixed microbial communities are viewed as one meta-organism. Metagenomics, metatranscriptomics,

metaproteomics and metametabolomics are employed to determine respectively the DNA sequences of the metaorganism under study, the collectively transcribed RNA, the translated proteins and the metabolites resulting from cellular processes. All of the generated data can then be used to identify the metabolic pathways and cellular processes at work within an ecosystem. Yet another level of information is required to access the molecular interactions occurring within the ecological niche under investigation, and this is achieved by the application of metainteractomics (Fig. 1; Lievens *et al.*, 2010; Medina & Sachs, 2010; Janga *et al.*, 2011). Ultimately, system approaches aim to develop mathematical models that can be used to predict the behaviour of a biological system in response to environmental stimuli (Fig. 1; Raes & Bork, 2008; Röling *et al.*, 2010).

Metaproteomics, which is the identification of all the proteins expressed at a given time within an ecosystem (as defined by Wilmes & Bond, 2004), is an indispensable element of system approaches and plays a key role in the determination of microbial functionality. Microbial metaproteomics has been applied in the context of diverse environments such as soil (Benndorf *et al.*, 2007; Williams *et al.*, 2010; Wang *et al.*, 2011), sediments (Benndorf *et al.*, 2009; Bruneel *et al.*, 2011), marine (Morris *et al.*, 2010; Sowell *et al.*, 2011), freshwater (Ng *et al.*, 2010; Habicht *et al.*, 2011; Lauro *et al.*, 2011), human intestinal tract (Verberkmoes *et al.*, 2009a; Rooijers *et al.*, 2011), human oral cavity (Rudney *et al.*, 2010), animal guts (Toyoda *et al.*, 2009; Burnum *et al.*, 2011) and natural and bioengineered systems (Ram *et al.*, 2005; Wilmes *et al.*, 2008a; Jehmlich *et al.*, 2010). Typically, metaproteomic approaches involve up to seven main steps (Fig. 2), namely sample collection, recovery of the targeted fraction, protein extraction, protein separation and/or fractionation, mass spectrometry analysis, databases searches and finally data interpretation, whereby the expressed proteins and pathways identified are used to access information about system functioning (for detailed descriptions of the methodologies involved, see Wilmes & Bond, 2006 and Verberkmoes *et al.*, 2009b). Each environment offers specific challenges and limitations within this workflow. Typically, sample collection and recovery of the targeted fraction (Fig. 2) can be problematic in the marine and freshwater context, where microorganisms can be recovered from hundreds of litres of water away from laboratory facilities. The protein extraction step (Fig. 2) has proven specifically difficult when dealing with soil samples, which naturally contain interfering humic acids. Such compounds are usually co-extracted together with proteins and are known to interfere with protein quantification, separation and identification (Bastida *et al.*, 2009). The use of gel-based methods for protein separation presents some disadvantages regardless of the origin of the

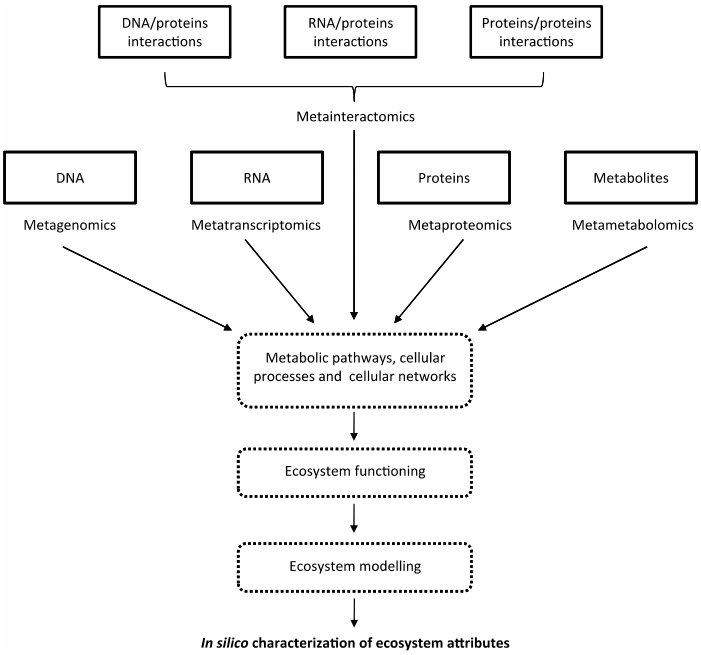


Fig. 1. System approach for the characterization of microbial ecosystems. Metagenomics (DNA sequencing of all the microorganisms from an ecosystem), metatranscriptomics (analysis of RNA collectively transcribed by all the microorganisms from an ecosystem), metaproteomics (analysis of proteins collectively expressed by all the microorganisms from an ecosystem) and metametabolomics (analysis of metabolites collectively produced by all the microorganisms from an ecosystem) are employed to access the metabolic pathways and cellular processes at work in an ecosystem. Metainteractomics (analysis of the molecular interactions between all the microorganisms from an ecosystem) is used to investigate the cellular network in an ecosystem. All the resulting data provide insights into ecosystem functioning and are used to generate a model, which in turn can allow the prediction of the behaviour of an ecosystem in response to environmental changes.

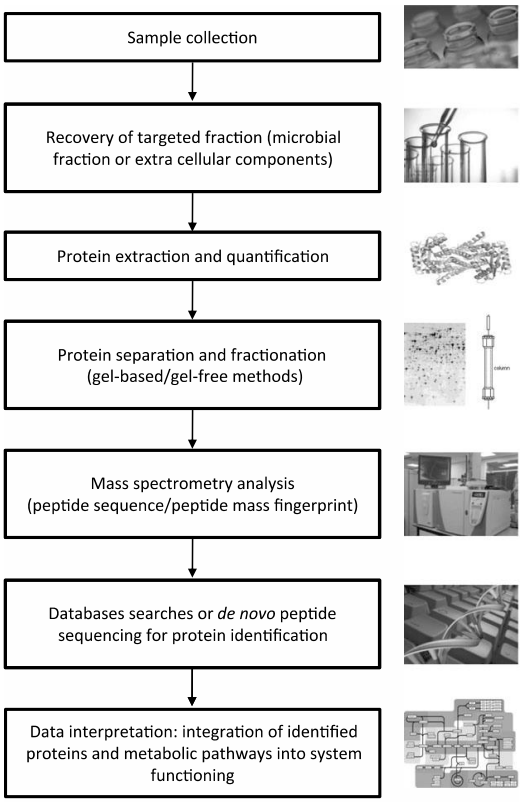


Fig. 2. Typical workflow for metaproteomics analysis.

sample. Such drawbacks are typically those associated with two-dimensional gel electrophoresis (2-DGE): proteins with extreme isoelectric points (basic or acidic) or extreme molecular weight (very large or very small), lipophilic proteins and low abundance proteins are typically excluded (Gygi *et al.*, 2000; Ong & Pandey, 2001). Finally, the limitations encountered in the last three steps of the metaproteomic workflow (Fig. 2), namely mass spectrometry analysis, databases searches and data interpretation are intrinsically linked to the success of the previous steps with a disadvantage in fields where no metagenome sequences are available. Overall, metaproteomics relies on the availability of relevant genome and metagenome sequences when searching generated mass spectra against existing databases for protein identification. As such, this approach cannot be viewed as an isolated method because it benefits from genome/metagenome sequencing for protein identification. However, when no relevant sequences are available, *de novo* peptide sequencing can be used for protein identification (Lacerda *et al.*, 2007; Fig. 2). In addition to being an integrative component of system approaches (Fig. 1), metaproteomics presents some valuable advantages over other 'omics' technologies for functional analyses. Primarily, metagenomic data only account for the microbial potential of a system and do not provide any insights into microbial activity. On the other hand, metatranscriptomics is one step closer to the identification of active metabolic pathways but does not allow for translational regulation to be taken into consideration; indeed, a lack of correlation between mRNA levels and proteins levels has been previously documented (Gygi *et al.*, 1999; Pradet-Balade *et al.*, 2001). Finally, metaproteomics provides significant insights into microbial activity together with metametabolomics, which is the study of the intermediate and end-products of cellular processes. Metagenomic data typically include numerous genes of unknown function (Ram *et al.*, 2005), most likely involved in novel functional systems. Metaproteomics may be useful to identify the circumstances under which these unknown functions are required and might therefore help to elucidate which systems hold the most potential for further investigation. Metaproteomics might also prove to be valuable for the identification of key microbial activities occurring in natural environments that could be exploited in a bioengineered context. For example, the investigation of the proteins expressed within the wood termite gut (Burnum *et al.*, 2011) or the sheep gut (Toyoda *et al.*, 2009) aimed at identifying natural microbial processes involved in the degradation of wood and cellulose. Such microbial processes could be harnessed for the production of renewable energy from wood or grass. For the purpose of this review, we discuss the advancement of metaproteomics in the context of human biology, soil, marine and freshwater environments as well as natural and bioengineered systems.

Protein expression in the human microbiome

In the context of human biology, the analysis of microbial community function has been investigated mainly in the intestinal tract and the oral cavity. Metaproteomic approaches in this field promise to be increasingly comprehensive because of the exponential progress in the generation of relevant genomic and metagenomic datasets. The Human Microbiome Project, which aims at characterizing the microbial communities present at several sites on the human body, has generated to date 178 full genome sequences from intestinal species (Nelson *et al.*, 2010), while gut metagenomic data were recently reported for 124 European individuals (Qin *et al.*, 2010). The reason behind the high activity in this field of research lies in the apparent links between microbial

communities and human health and disease. Applying metaproteomic approaches in such context has the potential to lead to the identification of protein markers that may be indicative of a healthy or a diseased state. In addition, the untargeted nature of metaproteomics makes it an ideal strategy to map unforeseen interactions between the human microbial communities and its host.

Despite the wealth of available genomic and metagenomic data, only a small number of metaproteomic studies have been reported to date. One of the first attempts at characterizing protein expression from the gut revealed the difficulty of such a task in the absence of relevant sequences in the databases. Klaassens *et al.* (2007) conducted a metaproteomic analysis of faecal samples from infants using two-dimensional gel electrophoresis (2-DGE) in combination with mass spectrometry analysis (matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, MALDI-TOF-MS). From 100 μg of protein, the authors could obtain more than 200 protein spots on each 2D-gel. Fifty-five spots were excised and analysed by mass spectrometry, but no protein identification could be obtained. After *de novo* sequencing was performed for four spots, 11 peptide sequences were determined, of which only one showed a high-level of similarity with a known protein (bifidobacterial transladolase). More recently, the identification of 2214 proteins from human faecal samples was reported using a shotgun metaproteomic approach (Verberkmoes *et al.*, 2009a). In this study, proteins were extracted from about 25 mg of microbial cells recovered from faecal material of twins, and the resulting peptide mixture was analysed using 2D-liquid chromatography-tandem mass spectrometry (2D- nano-LC-MS/MS). Mass spectra were searched against four databases, which all included two unmatched metagenome datasets (Gill *et al.*, 2006), and were supplemented with relevant sequences such as *Bacteroides* spp. genomes (Verberkmoes *et al.*, 2009a). One of the major breakthroughs from this study was the demonstration that unmatched metagenomes can be used to identify protein expressed from mixed microbial consortia (Verberkmoes *et al.*, 2009a). This finding was quite remarkable because each individual possesses a unique gut microbial communities (Zoetendal *et al.*, 2006). Another interesting result reported by Verberkmoes *et al.* (2009a) was the observation of a clear discrepancy in the distribution of clusters of orthologous groups (COG) categories between the metaproteome and the metagenome. This emphasizes the advantages of using a metaproteomic approach, whereby the measure of expression and translation of a gene product can be achieved, over a metagenomic approach, which is only indicative of the potential of a gene to be expressed and translated. More recently, an iterative workflow has been successfully developed to achieve an optimized use of the continuously growing genomic and metagenomic databases and has led to the identification of up to 5000 peptides from one sample (Rooijers *et al.*, 2011). In this study, proteins were extracted from the microbial fractions recovered from faecal samples of two individuals. The extracted proteins (50 μg per sample) were separated by one-dimensional polyacrylamide gel electrophoresis (1D-PAGE) prior to LC-MS/MS analysis. Metagenomic data matched to these two faecal samples were also collected. Metagenome annotation has been identified as the main shortcoming for the utilization of such data for protein identification (Keller & Hettich, 2009; Verberkmoes *et al.*, 2009b; Röling *et al.*, 2010). The strength of the iterative workflow proposed by Rooijers *et al.* (2011) is the use of nonannotated, unassembled metagenome sequences in conjunction with robustly annotated genomes from single microorganism. This iterative workflow could be systematically applied to further metaproteomic analyses provided that metagenome sequences representative of the system under investigation are available.

The characterization of the microbial communities from the human oral cavity is also gaining momentum with the Human Oral Microbiome Project, which has made available an extensive 16S rRNA gene sequence database from oral microorganisms and with increasing efforts to determine the genetic diversity of oral microbial communities (Lazarevic *et al.*, 2009; Dewhirst *et al.*, 2010; Crielaard *et al.*, 2011). The oral microbiota has been associated with infectious diseases such as periodontitis and tonsillitis, but also with systemic diseases such as stroke and pneumonia (Joshipura *et al.*, 2003; Awano *et al.*, 2008; Dewhirst *et al.*, 2010). The microbial community of the oral cavity has also been suggested as a diagnostic marker for oral cancer, specifically targeting the abundance of three bacterial species; *Capnocytophaga gingivalis*, *Prevotella melaninogenica* and *Streptococcus mitis* (Mager *et al.*, 2005). However, up until now, the characterization of the microbial communities from the human mouth has been mainly directed at determining the diversity more than the functionality of such communities. Indeed, when metaproteomic data are generated for oral cavity samples, most studies so far only reported on the resulting phylogeny obtained from peptide species assignment and did not investigate the functions of the proteins identified (Guo *et al.*, 2006; Xie *et al.*, 2008; Grant *et al.*, 2010). Realizing the potential of the generated data, some of the authors involved in the study of Xie *et al.* (2008) revisited their results and provided a thorough analysis of the functions represented by the proteins expressed in the oral cavity (Rudney *et al.*, 2010). In that respect, this study constitutes the only true metaproteomic analysis conducted on the human oral microbiota. The authors

successfully developed a three-step method to effectively characterize the proteins expressed in salivary samples and identified over 1000 human proteins and 139 microbial proteins from 200 μg of protein extract (Xie *et al.*, 2008; Rudney *et al.*, 2010). The oral microbial mixed community was found to be metabolically active and mainly involved in the transport of carbohydrate. One of the major hurdles however in this field of research is the absence to date of oral metagenomic data. Once such data become available, future comparative metaproteomic studies should aim at characterizing community function in response to disease to identify functional biomarkers that could be used to develop diagnostic tests.

Protein expression in soil

Soil is widely regarded as a difficult medium for the analysis of microbial community structure and function. Although soil contains an abundance of microorganisms (~10^9 cells g^{-1} soil; Rosselló-Mora & Amann, 2001), the high species diversity, particularly of uncultured organisms, has hindered analytical progress. In addition, the presence of inhibitory substances, such as humic acids, has proved difficult for the extraction of nucleic acids and proteins from soil, without compromising downstream processing methods, or resulting in protein degradation (Solaiman *et al.*, 2007). As a result of these issues, the success of soil metaproteomics has been somewhat limited. Initial studies focused on investigating the optimization of protein extraction protocols that remove interfering compounds without inhibition of successive analysis (for details, see review from Bastida *et al.*, 2009; Chen *et al.*, 2009). One of the earliest reports of attempted identification of soil proteins from excised 1D-PAGE bands was carried out by Schulze *et al.* (2005). This study investigated the extracellular protein fraction from soil microparticles using LC-MS/MS analysis. The resulting mass spectra were searched against the NCBI database, which led to the identification of 75 proteins involved in the degradation of biological matter and included cellulases, collagenases, lignin phenoloxidases and laccases. Fifty per cent of the identified proteins were assigned to bacterial species, while the remaining proteins originated mainly from plants, fungi, vertebrates and nematodes. An interesting observation from this study was the redundancy of protein function as similar proteins were expressed by a range of species. For example, cellulases were expressed by three different bacteria and one fungus. More recently, a direct protein extraction method led to the identification of 716 proteins from 5 g of soil (Chourey *et al.*, 2010). In this study, the protein fractions were analysed using 2D-nano-LC-MS/MS, and the databases used for protein identification consisted of all the fully sequenced microbial genomes from the Integrated Microbial Genomes database in addition to an unmatched metagenome dataset from agricultural soil (Tringe *et al.*, 2005). The metaproteome was found to reflect survival strategies such as bacterial spore formation, probably to overcome the extended summer drought typical of the soil under investigation (Chourey *et al.*, 2010). Another study (Wang *et al.*, 2011) individually analysed the rhizospheric soil metaproteome associated with rice, sugar cane and the flowering plant *Rehmanniae*. Separation of extracted proteins by 2-DGE allowed the visualization of approximately 1000 individual proteins on each gel, with 538 common to all three samples, from which a subset of 287 spots were randomly selected for protein identification. Using MALDI-TOF/TOF followed by MS/MS and the NCBI database to search the resulting mass spectra, 189 spots were successfully identified, including 107 proteins attributed to plants, 72 to bacteria and fungi and 10 to fauna. The identified proteins were associated with a wide range of functional categories including, amongst others: energy metabolism, protein turnover, amino acid biosynthesis and proteins involved in resistance mechanisms.

In the soil environment, metaproteomics can be used to understand complex community interactions associated with *in situ* bioremediation of contaminated soil sites. Earlier studies such as Renella *et al.* (2002) and Singleton *et al.* (2003) focused on metal contaminants such as cadmium, while more recent studies moved towards organic contaminants such as 2,4-dichlorophenoxy acetic acid (Benndorf *et al.*, 2007), toluene (Williams *et al.*, 2010) and diesel fuel (Bastida *et al.*, 2010). Benndorf *et al.* (2007) conducted metaproteomic analyses on soil microcosms to investigate the degradation of 2,4-dichlorophenoxy acetic acid (2,4-D). Soils were percolated with 2,4-D solution to promote the development of indigenous degrading species, and one of the microcosms was also augmented with a mixture of bacteria known to degrade 2,4-D. Proteins were extracted from 5 g of soil and separated by 1D-PAGE, prior to excision of protein bands. Protein identification was carried out using nano-LC-electrospray ionization (ESI) source-MS/MS, and searching the resulting mass spectra against the NCBI database. A total of only four proteins were identified, and the low level of protein identification was attributed to the difficulties encountered during soil protein extraction because of the presence of interfering humic compounds. Despite the limited number of proteins identified, the authors could conclude that the autochthonous soil community showed similar 2,4-D degradation capability as that of the augmented soil sample based on the function of the identified soil proteins (specifically, 2,4-dichlorophenoxyacetate dioxygenase), as well as the monitoring of 2,4-D

degradation using HPLC separation combined with UV detection (Benndorf *et al.*, 2007). More recently, changes in the metaproteome of a soil microbial community were investigated in response to glucose and toluene amendment (Williams *et al.*, 2010). The microbial fraction was isolated from 5 g of soil, and the corresponding extracted proteins were separated by 1D-PAGE and analysed by MALDI-TOF/TOF followed by MS/MS. This study detected 187 proteins in total amongst which 47 were identified when searching mass spectra against the NCBI database. Proteins specifically associated with known toluene degradation pathways could not be identified, but the metaproteomic data overall reflected microbial adaptation to the presence of toluene, noted by the expression of stress-related proteins (Williams *et al.*, 2010). One of the major limitations of this study, however, is the lack of measurement of the level of toluene degradation, which consequently does not allow for any conclusions regarding the toluene degrading capabilities of the soil community.

Also focusing on bioremediation, Bastida *et al.* (2010) compared the protein content of an artificially hydrocarbon-amended soil with an unpolluted control. Proteins were extracted from the microbial fraction recovered from 10 g of soil and either separated by 1D-PAGE followed by LC-ESI-MS, or by direct LC-ESI-MS. Mass spectra were searched against the bacterial entries of the NCBI database. The majority of the 42 identified proteins were related to cellular metabolism, indicating high microbial activity both in the control and the polluted soil; however, no specific known hydrocarbon degradation enzyme could be detected in the polluted soil (Bastida *et al.*, 2010). Once again in this study, no measurement of the level of pollutant degradation was reported, and therefore, no information regarding the degrading capabilities of the soil communities was provided.

Overall, the field of soil metaproteomics is still in its infancy, mainly because of the technical difficulties typical of this environment. Improvements in extraction protocols are reported regularly, and even though only one soil metagenomic study has been published to date (Tringe *et al.*, 2005), it has been largely underused. Strikingly, the highest number of proteins identified by soil metaproteomics so far (716 proteins) was based on the use of databases including an unmatched soil metagenome (Chourey *et al.*, 2010). This study however was still a proof of concept, where the authors aimed to demonstrate the technical feasibility of soil metaproteomics and did not really discuss identified protein functions.

An interesting alternative strategy to metaproteomics, which has been employed in soil, is the use of functional metagenomics, where metagenomic DNA is expressed in a surrogate host, typically *Escherichia coli*. The resulting clones are subsequently screened for specific activities such as antibiotic resistance (Allen *et al.*, 2009; Donato *et al.*, 2010). Functional metagenomics has proven very useful for the identification of novel enzymes involved in antibiotic resistance and can be used as a targeted strategy to uncover specific metabolic activities. However, functional metagenomics does not provide insights into *in situ* processes, but reveals the function of metagenomic DNA in *E. coli*.

Future studies in this field should aim at combining metagenomics with metaproteomics within the soil environment, which is largely dominated by uncultured microorganisms. Taking advantage of the only soil metagenome available to date (Tringe *et al.*, 2005), attempts should be made to reanalyse existing mass spectra using databases inclusive of this metagenome. Furthermore, the pending availability of soil metagenomic data as instigated by the TerraGenome Project (Vogel *et al.*, 2009) should considerably improve the success of protein identification from soil environmental samples. In addition, when investigating contaminant degradation, efforts must be made to ensure that pollutants levels are monitored throughout the duration of the experiment to clarify whether the nonidentification of proteins associated with contaminant degradation is because of the lack of degradation activity or is masked by more abundant processes. Remarkably, despite the low number of identified proteins, Benndorf *et al.* (2007) could correlate the measured contaminant degradation with the presence of specific pollutant degrading proteins in their soil samples. This might imply that other bioremediation studies (Bastida *et al.*, 2010; Williams *et al.*, 2010), which did not include the monitoring of the level of contaminants in their experimental setup, might have been investigating pollutants at concentrations that did not allow for the soil community to utilize its degrading capabilities.

Protein expression in marine and freshwater

Aquatic regions of the world harbour significant microbial populations with an average of ~2.5×10^6 cells ml^{-1} of seawater (Kan *et al.*, 2005). These natural mixed microbial communities play crucial roles in essential processes, such as carbon and nitrogen cycling and organic matter decomposition (Maron *et al.*, 2006). Metagenomic approaches have revealed the extensive microbial diversity and metabolic potential of marine and freshwater mixed communities (Venter *et al.*, 2004; Rusch *et al.*, 2007; Ng *et al.*, 2010). These metagenomics datasets provide an ideal platform for metaproteomics, which is a growing field both in marine and freshwater ecosystems.

Metaproteomics aims at capturing the protein expression profile at the point of sampling, and thus, limiting interferences with the natural microbial consortia in samples is of vital importance. Typically, once water samples are collected, the microbial fraction is immediately concentrated prior to storage pending protein extraction. To stop further protein expression in response to sample handling conditions, strategies such as chemical fixation (Kan *et al.*, 2005) or microwave fixation (Mary *et al.*, 2010) have been employed. The majority of studies, however, transfer concentrated pellets directly to storage at −80 °C or employ liquid nitrogen flash freezing (Sowell *et al.*, 2009, 2011; Morris *et al.*, 2010).

As discussed earlier in the human biology context and in the soil environment, identification of the protein expressed by mixed microbial consortia has proven to be challenging without the availability of relevant metagenomic data. Kan *et al.* (2005) conducted the first metaproteomic analysis of a complex aquatic microbial community from Chesapeake Bay using 2-DGE in combination with both MALDI-TOF and LC-MS/MS. From 100 μg of extracted proteins, the authors could detect over 200 protein spots on each 2D-gel. Forty-eight spots were excised, from which no protein identification could be obtained using MALDI-TOF analyses. Seven of the 48 spots were further analysed using LC-MS/MS and *de novo* sequencing, which led to the identification of three proteins that shared homologies with proteins annotated from the Sargasso Sea metagenome (Venter *et al.*, 2004). More recently, a metaproteomic study investigated the functionality of specific groups of microorganisms known to be abundant in the oligotrophic Sargasso Sea (Sowell *et al.*, 2009). Protein extracts from microbial cells recovered from two samples of ~230 L of surface seawater were analysed using LC-MS/MS. The resulting mass spectra were searched against three synthetic metaproteomes comprising, respectively, the SAR11 clade, *Prochlorococcus* and *Synechococcus* genomic fragments from the Sargasso Sea metagenome (Venter *et al.*, 2004). This approach led to the identification of 236 SAR11 proteins, 402 *Prochlorococcus* proteins and 404 *Synechococcus* proteins, which were found to largely reflect cellular adaptations to stringent environmental conditions under which the microorganisms are competing for nutrients (Sowell *et al.*, 2009). A comparative metaproteomic study investigated the membrane proteins expressed by the mixed microbial communities from surface waters in the South Atlantic (Morris *et al.*, 2010). This study compared metaproteomic profiles over a large geospatial natural gradient extending from an oligotrophic gyre to a nutrient-rich coastal upwelling region. Microorganisms were recovered from 10 samples of 100–200 L each, and the enriched extracted membrane protein fractions were analysed by LC-MS/MS. The resulting mass spectra were searched against an extensive database from the Global Ocean Sampling (GOS) metagenomic project containing over 600 000 predicted proteins (Rusch *et al.*, 2007; Yooseph *et al.*, 2007). In total, this study identified 2273 proteins with 428 ± 158 distinct membrane proteins per sample (Morris *et al.*, 2010), compared with 1042 proteins identified when whole protein extracts were analysed (Sowell *et al.*, 2009). To illustrate the need for relevant metagenomic sequences when conducting metaproteomics analyses, Morris *et al.* (2010) searched their mass spectra against the GenBank database and recorded a staggering 6.2-fold decrease in peptide identification compared with searching against the extensive Global Ocean Sampling database. Overall, this study reported different metabolic activities as indicated by protein expression along the natural nutrient gradient investigated. Specifically, Ton-B transport systems, known to use a proton motive force for nutrient membrane translocation, were enriched in the nutrient-rich coastal samples, while porins and permeases were preferentially expressed in the oligotrophic open ocean samples. In addition, urea ABC transporters and photosystem proteins were more abundant in the open ocean, when compared with the coastal area. Overall, the authors concluded that the membrane metaproteomic data and derived species assignment reflected well the microbial biodiversity and physicochemical characteristics of the oligotrophic open ocean and nutrient-rich environments investigated (Morris *et al.*, 2010).

Another study employed metaproteomics to investigate the functionality of the marine microbial communities in the nutrient-rich Oregon coastal seawater upwelling region (Sowell *et al.*, 2011). The marine microorganisms were isolated from ~100 L of surface seawater, and the extracted whole protein fractions were analysed by 2D-nano-LC-MS/MS. The generated mass spectra were searched against a database composed of the predicted proteins from the GOS project (Rusch *et al.*, 2007) that was edited to contain only the sequences found in similar natural environments, as well as additional genome sequences from two Oregon coastal isolates (Sowell *et al.*, 2011). Overall, this study identified 481 unique protein families, with 29% of the total detected spectra related to transport functions. The types of transport proteins identified in this nutrient-rich environment (e.g. specific for nitrogen, carbon and sulphur containing compounds) differed from those detected in the oligotrophic Sargasso Sea (e.g. specific for phosphate and phosphonate, known to be abundant during phosphorus starvation; Sowell *et al.*, 2009). Differential protein expression as a response to natural nutrient gradients within the marine environment was also observed in the South Atlantic ocean (Morris *et al.*, 2010).

In freshwater environments, two recent studies investigating microbial community functions in Ace Lake Antarctica directly combined metagenomics with metaproteomics (Ng *et al.*, 2010; Lauro *et al.*, 2011). Ng *et al.* (2010) targeted a dominant green sulphur bacterium (*Chlorobioaceae*) known to be prevalent at specific depths (12–14 m) in the water column, while Lauro *et al.* (2011) investigated the microbial communities at work throughout the water column (5–23 m). In both studies, microbial fractions were recovered from 1- and 10-L samples obtained after drilling the ice cover of the lake. Whole protein extracts were analysed using 1D-PAGE followed by LC-MS/MS. Mass spectra were directly searched against the assembled metagenome from the same sample (Ng *et al.*, 2010; Lauro *et al.*, 2011).

Metagenomic analysis of samples from 12.7 m depth of Ace Lake assigned 76% of the predicted open reading frames (ORF) to a single species of green sulphur bacterium, (*C*-Ace; Ng *et al.*, 2010), highlighting the predominance of this organism in this environment. The resulting composite genome was then directly used to construct a database for searching metaproteomic mass spectra to inform on the functional activities of *C*-Ace at the time of sampling. This approach led to the identification of 504 proteins corresponding to ~31% of the total predicted proteome of *C*-Ace (Ng *et al.*, 2010). As observed in the context of the gut microbiome (Verberkmoes *et al.*, 2009a), the comparison of the distribution of COG categories between the metaproteome and the metagenome revealed that some functional categories were misrepresented in the metaproteome. Specifically, proteins involved in translation and energy production and conversion were statistically overrepresented, while amongst others, proteins involved in defence mechanisms and inorganic ion transport and metabolism were underrepresented. In addition, the analysis of genes and corresponding proteins that did not have any orthologues in known green sulphur bacteria gave some insights about potential adaptive mechanisms in *C*-Ace. Evidence of a requirement for specific polysaccharides structures was attributed to cold adaptation, while the presence of genes encoding DNA restriction and modification system could present a protective advantage against bacteriophages. Also worth noting was the construction of a pathway for sulphide oxidation for *C*-Ace, and the lack of evidence for assimilatory sulphate reduction suggesting the strict dependence of *C*-Ace on sulphate-reducing bacteria, located at 14 m depth in Ace Lake (Ng *et al.*, 2010). To conclude, the authors could access the proteins necessary for the success of *C*-Ace in Ace Lake under cold, oligotrophic, oxygen limited and extreme light conditions and therefore access the biology of this organism in its natural habitat (Ng *et al.*, 2010). The structure and functions of the microbial communities in Ace Lake were further investigated throughout the water column by sampling from six depths with the aim of capturing the interactions between microbial populations that defined nutrient cycling (Lauro *et al.*, 2011). The metagenomic analysis could assign ~28% of open reading frames to COG categories, while the metaproteomic study identified 1824 proteins. Interestingly, phylogenetic diversity was found to increase with depth and was associated with an increased amount of hypothetical proteins, which accounted for 67% of the identified proteins at 23 m depth. In addition, most of these proteins did not match any orthologues from known microorganisms and could belong to novel functional pathways (Lauro *et al.*, 2011). Combining the physicochemical data available for Ace Lake with matching metagenomic and metaproteomic data, Lauro *et al.* (2011) could describe the carbon, nitrogen and sulphur cycles throughout the water column. Briefly, for the carbon cycling, cyanobacteria were found to carry out aerobic carbon fixation in the upper stratum of the lake, anaerobic carbon fixation occurred further down the water column and was performed by green sulphur bacteria conjointly with sulphate-reducing bacteria, while fermentation processes were thought to take place at the bottom of the lake. In addition, remineralization of particulate organic carbon to dissolved organic carbon was found to occur at the surface of the lake with further heterotrophic conversion carried out by *Actinobacteria* and members of SAR11 clade. In the lower stratum of the lake, remineralization of particulate organic carbon was thought to result from the joint activities of fermentative, sulphate-reducing and methanogenic microorganisms ultimately leading to the production of CO_2 and CH_4. Carbon monoxide oxidation was also thought to be an important energy generation pathway throughout the water column as indicated by the detection of CO dehydrogenase genes (Lauro *et al.*, 2011). The nitrogen cycle was found to involve nitrogen assimilation throughout the lake, as indicated by the detection of glutamine and glutamate synthetases in the metaproteome, with remineralization localized at the lower stratum. Interestingly, no evidence of nitrification occurring throughout the water column could be detected. Indeed, the metagenome did not include any ammonia oxidation genes or any signatures of known nitrifying bacteria or archaea. As Ace Lake is associated with continuous low level of nitrate, the authors suggested that the absence of nitrification might be a strategy to conserve bioavailable nitrogen (Lauro *et al.*, 2011). Finally for the sulphur cycle, green sulphur bacteria were found to consume the sulphide produced by the sulphate-reducing bacteria, to generate sulphate and thus providing the replenishment of the sulphate-reducing bacteria substrate and ensuring the continuous

turnover of sulphur compounds in the lake. Even though genes for assimilatory sulphate reduction were detected in the metagenome, no evidence of their expression could be found in the metaproteome. However, dissimilatory sulphide reduction was found to be an active pathway as indicated by the identification of proteins from sulphide reductase complex from green sulphur bacteria (Lauro *et al.*, 2011). Because of the crucial role seemingly played by *C*-Ace in the lake, as outlined by metagenomics and metaproteomics, and the low number of viral particles where this bacterium is located (12.7 m depth), the authors suspected that *C*-Ace was not following the previously developed model of viral/population dynamics in aquatic environments (Rodriguez-Brito *et al.*, 2010). The authors attributed this unusual observation to the influence of extreme light conditions on the microorganism biology. To test their hypothesis, a mathematical model was developed, where the growth of *C*-Ace was intrinsically linked to the light intensity in the lake (Lauro *et al.*, 2011). From this, the authors could predict that the persistence of *C*-Ace in the lake was because of the absence of phage. By incorporating this prediction with the information deduced from metagenomics and metaproteomics, it was concluded that the emergence of phage predators could deplete the Ace Lake of green sulphur bacterial populations, which, because of their central role in the recycling of carbon, nitrogen and sulphur, would have severe consequences for the whole lake community (Lauro *et al.*, 2011). Overall, these two studies (Ng *et al.*, 2010; Lauro *et al.*, 2011) focused on the interpretation of the data generated using metagenomics and metaproteomics and are quite unique in that aspect as they did not stop at the demonstration of the feasibility of applying such technologies to Ace Lake but fully used their results to gain insights into the biology of this ecosystem.

In most environments, the difficulties associated with attempting to characterize the metaproteome without matched or unmatched metagenome sequences have been demonstrated. However, this does not apply when investigating natural environments largely dominated by one species, where protein identification can be obtained using a single sequenced relevant microbial isolate (Habicht *et al.*, 2011). Conversely, one could argue that these types of study do not analyse the metaproteome but simply investigate the microbial behaviour of one species within its natural environment. Such a study identified 1321 proteins from *Chlorobium clathratiforme* (species representing more than 50% of the bacterial population) throughout the water column in Lake Cadagno, Switzerland (Habicht *et al.*, 2011). Interestingly, when searching the mass spectra generated by LC-MS/MS against all *Chlorobi* genomes, an additional 350 proteins could be identified, revealing the presence of species not previously detected in the lake. A search against the entire UniProt database led to the identification of a further 120 proteins from various species. Such a study demonstrates that, if the microbial communities investigated are largely dominated by known species, LC-MS/MS analysis can lead to extensive results in the absence of relevant metagenomic data. However, even if such results are valid, they are nonetheless biased towards dominant species and might not be suitable for integration in system approaches (Fig. 1).

Protein expression in natural and bioengineered systems

Scientists have long known the significance of the role played in the environment by mixed microbial communities. However, the exact processes carried out by these natural consortia are far from resolved. Even in relatively simple natural biofilms, the expression of thousands of proteins is required to enable survival and growth, including proteins involved in nutrient metabolism, stress response and environmental signalling (Ram *et al.*, 2005). One such biofilm community, originating from an acid mine drainage system within the Richmond mine at Iron Mountain, California, has been extensively studied (Ram *et al.*, 2005; Lo *et al.*, 2007; Belnap *et al.*, 2010; Denef *et al.*, 2010; Mueller *et al.*, 2010, 2011; Wilmes *et al.*, 2010; Pan *et al.*, 2011). Indeed, one of the first publications investigating mixed microbial consortia from this site, Ram *et al.* (2005), is widely considered as a breakthrough study in environmental metaproteomics. Even though very significant, most of these studies however (Lo *et al.*, 2007; Denef *et al.*, 2010; Wilmes *et al.*, 2010; Mueller *et al.*, 2011) do not really fall under the umbrella of metaproteomics, but as in the case of Habicht *et al.* (2011), they investigate the physiology of one dominant species within its natural habitat. Nonetheless, these approaches collectively participated in important methodological advances and will be briefly discussed here.

Lo *et al.* (2007) demonstrated the use of proteomics to determine genomic divergences between uncultivated microorganisms and closely related sequenced species within the acid mine drainage biofilm. Denef *et al.* (2010) employed this strategy to differentiate between two closely related genotypic groups of *Leptospirillum* group II (99.7% similarity between 16S rRNA gene sequences). While one genotypic group typically prevails in early-stage biofilm, the other is dominant in the later stages of biofilm formation. To investigate this ecological divergence, proteomic data previously published from 27 samples (Denef *et al.*, 2009) were analysed for the determination of ecologically preferential functions. The authors identified a subset of proteins that could be

associated with the observed ecological divergence. Specifically, proteins originating from an early-stage biofilm were found to be mainly involved in co-factor biosynthesis and motility, while proteins associated with energy production and conversion emerged as biofilm maturity progressed (Denef *et al.*, 2010). Applying a similar strategy, Wilmes *et al.* (2010) focused on the genomic divergence of two dominant species within the acid mine drainage biofilm, namely *Leptospirillum* groups II and III, using a combination of proteomics and metabolomics. A total of 765 proteins and 3740 metabolites were detected across 14 biofilm samples. Each of the two *Leptospirillum* groups investigated could be associated with a distinct cluster of proteins and metabolites, indicating genomic divergence between the two bacteria. From their data, the authors could identify a limited niche overlap and a low level of interspecies competition, suggesting separate resource utilization for each species within the biofilm (Wilmes *et al.*, 2010).

Investigating further the acid mine drainage biofilm, Mueller *et al.* (2010) combined metaproteomics with geochemical and biological data to investigate the behaviour of the biofilm microbial communities along environmental gradients. The analysis of 28 biofilm samples led to the identification of a total of 6296 proteins and differences in protein expression was found to correlate with environmental parameters (Mueller *et al.*, 2010). For example, temperature at the time of sampling was found to impact on levels of cold-shock proteins and proteins involved in fatty acid biosynthesis, expressed at a lower level and at a higher level with increasing temperature, respectively. In addition, a shift in the metaproteome was observed from ribosome biosynthesis and transcription in the early stage of biofilm formation to environmental signalling, chemotaxis and biosynthesis of extracellular components in the mature biofilm (Mueller *et al.*, 2010). In a subsequent study, Mueller *et al.* (2011) focused their investigations on the initial dominant species, *Leptospirillum* Group II, and its evolving protein expression throughout the development of the biofilm community. Using replicate samples and analysing membrane and cytoplasmic protein fractions, this study aimed at identifying key proteins that could be used as biomarkers to estimate biofilm maturity (Mueller *et al.*, 2011). The authors identified a total of 4107 distinct proteins, corresponding to ~45% of the predicted proteome of *Leptospirillum* Group II. Changing protein abundance patterns were successfully determined as a function of biofilm maturity. For example, the metabolism of simple carbon compounds was found to dominate early growth stage biofilms, while the metabolism of complex carbohydrates increased in late stages of biofilm formation (Mueller *et al.*, 2011). Although Mueller *et al.* (2011) briefly mentioned the use of ^{15}N for quantification of expressed protein, an earlier study by Belnap *et al.* (2010) investigated quantitative proteomic analysis more thoroughly to compare a laboratory-grown biofilm with a natural biofilm. The aim of this study was to validate the use of laboratory model systems for the investigation of ecological hypotheses. In total, the relative quantification of over 2500 proteins between natural- and laboratory-grown biofilms was determined. This comparison showed that improvements in laboratory culturing conditions resulted in the decrease of metabolic stress protein expression, such as proteins involved in defence mechanisms and oxidative damage repair, and this correlated with increased community growth rates (Belnap *et al.*, 2010). This study highlighted the relevance of using model systems under laboratory controlled conditions to gain some understanding of natural microbial processes. A subsequent study by Pan *et al.* (2011) used this laboratory culture model to successfully develop a high-throughput stable isotope probing (SIP) method. Employing this new technique, the authors could trace the incorporation of ^{15}N into thousands of proteins expressed by a mixed microbial community. Also applying SIP technology, Bastida *et al.* (2011) quantitatively traced the movement of ^{13}C-benzene between trophic levels, specifically from bacteria to eukaryotic organisms, such as *Chironomus* sp. larvae. SIP-proteomics methods could prove very valuable to track down fluxes of ^{13}C or ^{15}N in mixed microbial communities. However, such setups could not be applied *in situ* and are limited to laboratory-based models including microcosms.

Although the incorporation of labelled carbon and nitrogen has been investigated in the context of laboratory-grown mixed anoxic communities (Jehmlich *et al.*, 2008), to the best of our knowledge, no protein-based SIP analysis of complex, engineered systems has been yet undertaken, which could allow for optimization of economically significant processes. However, metaproteomic studies of engineered processes, such as activated sludge systems, are extensive (Wilmes & Bond, 2006; Park & Novak, 2007; Park *et al.*, 2008a, b; Abram *et al.*, 2011). Activated sludge systems are principally employed for the treatment of wastewater in developed countries. The aggregation of microbial biomass to form well-settling flocs has been identified as a key parameter for the successful operation of wastewater treatment plants. The study of the metaproteome of laboratory-scale-activated sludge systems has been divided into two approaches: analysis of the extracellular polymeric substances (EPS) that bind the microbial cells together, and analysis of the protein fraction located within the microbial cells themselves.

Researchers from the University of Massachussetts/Virginia Tech focused on the extraction and analysis of the

protein component of EPS. Their first study (Park & Novak, 2007) identified that EPS bound to different cations require different extraction methods, which resulted in the extraction of 'sub-proteome' components of the activated sludge sample. Focusing on the requirement for the reduction of volatile solids to 38% prior to land spreading of digested sludge (US EPA, 2003), Park *et al.* (2008a) demonstrated the selective degradation of divalent cation-bound proteins and aluminium-bound proteins by aerobic digestion, but not by the anaerobic treatment process, while the opposite was true of organic matter bound with ferric iron. This is of significant industrial importance, as it specifically indicates by protein analysis, that a two-step treatment process, featuring both anaerobic and aerobic stages, may be required for enhanced sludge treatment.

Next, Park *et al.* (2008b) attempted to identify proteins that could serve as biomarkers to monitor facility operations in activated sludge processes. Eleven intense bands from 1D-PAGE were investigated using LC-MS/MS and the whole NCBI database. Seven of these bands contained identifiable proteins associated with bacterial defence, cell appendages, cell surface outer membrane proteins, intracellular materials and influent sewage proteins. Interestingly, of the bands that did not return any positive matches, identifiable amino acid fragmentation patterns were obtained, indicating that these proteins are likely to originate from unsequenced microorganisms in full-scale-activated sludges, again indicating the requirement for metagenomic sequence data. As seen in other environments (human microbiome, soil, marine and freshwater), the use of unmatched metagenome sequences may have improved the success of protein identification in the above investigation. A short follow-up study analysed one of the samples from Park *et al.* (2008b) over the course of a 30-day batch digestion (Park & Helm, 2008). Three 1D-PAGE bands, additional to those studied in Park *et al.* (2008b), were investigated by LC-MS/MS using the whole NCBI database. All three bands (1) were associated only with divalent cations, (2) emerged during the batch anaerobic digestion of the activated sludge and (3) were identified as different subunits of methyl-coenzyme M reductase, a key enzyme for methane formation in methanogenic archaea, and in this case associated with *Methanosarcina barkeri*.

The intracellular microbial protein fraction of activated sludge has been investigated by a separate group of researchers. Wilmes & Bond (2006) applied metaproteomics to activated sludge originating from the end of the anaerobic and aerobic cycles of the same sequencing batch reactor (SBR) during stages of enhanced biological phosphorus removal (EBPR) and non-EBPR (nEBPR). From 100-mL sludge samples, separation by 2-DGE allowed the detection of an average of 665 spots per gel, although spot patterns differed significantly between EBPR and nEBPR samples. Furthermore, proteins patterns detected from the EBPR sludge samples were more consistent between aerobic and anaerobic stages, indicating the development of a more stable microbial community, which was not noted in the nEBPR samples. Although no attempt at protein identification was undertaken, the importance of this study must not be underestimated, as the demonstration of a statistically different protein expression between a functioning and a nonfunctioning system justifies the investigation of the metaproteome for identification of performance issues associated with full-scale-activated sludge processes.

A subsequent study by this group focused on the identification of proteins separated by 2-DGE. Using a combination of MALDI-TOF/MS and quadrupole-TOF MS/MS, and searching against three unmatched EBPR metagenomic databases, Wilmes *et al.* (2008a) identified 38 of 111 excised protein spots, corresponding to 33 unique proteins. It is suggested that the low identification rate (41%) may be due to strain variation between the sample sludge (from a UK SBR) and the three metagenomic databases employed (two from the United States and one from Australia). Possibly recognizing the technical limitations of 2-DGE, Wilmes *et al.* (2008b) then carried out analysis of extracted proteins directly via 2D nano-LC, followed by MS/MS analysis. Peptide tandem mass spectra were matched *in silico* to the three separate sludge metagenomic databases employed in Wilmes *et al.* (2008a), and against a combined database of the three metagenomic datasets. This extensive analysis resulted in the identification of a total of 5029 proteins, with 36% associated with *Accumulibacter phosphatis*, a major polyphosphate accumulating organism.

While activated sludge technology is commonly employed in the field of sewage treatment, anaerobic digestion technology is becoming increasingly popular for the treatment of recalcitrant chemical-containing wastewater streams. Jehmlich *et al.* (2010) were one of the first reported studies to investigate the metaproteome of an anaerobic community, originating from a batch sulphate-reducing enrichment culture, exposed to toluene. This study analysed 150 μg of protein by 2-DGE and identified 202 proteins from 236 excised gel spots using nano-LC-ESI-MS/MS analysis and searching the resulting mass spectra against the bacterial entries of the NCBI database. Identified proteins were found to be associated with both sulphate reduction and toluene degradation. Recently, the first metaproteomic investigation of granular biomass originating from a continuous anaerobic bioreactor was described (Abram *et al.*, 2011). The extracted proteins were separated by 2-DGE, which resulted in the detection

of 388 reproducible protein spots, of which 70 were excised. Thirty-three proteins were positively identified using nano-LC-ESI-MS/MS and searching the resulting mass spectra against both the whole NCBI database and the Trembl database. Proteins associated with the production of methane and the degradation of glucose were identified. One of the major hurdles in the field of anaerobic community function is the lack of metagenomic data, which would facilitate the identification of proteins.

Conclusions

Overall, the field of metaproteomics is gaining momentum at an exponential rate within very diverse environments. An overview of selected studies from the ecosystems discussed in this review is shown in Table 1. Advances in metaproteomics finally allow for the consideration of the integration of such data in system approaches (Fig. 1). This was partly achieved in the aquatic environment where Lauro *et al.* (2011) combined metagenomic, metaproteomic and physicochemical data to describe the interaction between the microbial populations defining the biogeochemical cycles throughout a water column. Such an approach could feasibly be transferred to other environmental ecosystems. To date, the application of complex system approaches is still scarce and requires a coordinated experimental design that brings together expertise from each of the many technologies involved.

The technical limitations encountered throughout the metaproteomic workflow (Fig. 2) have, for the most part, been addressed in the ecosystems discussed in this review. However, it should be kept in mind that an exhaustive investigation of the entire metaproteome is unlikely due to the unfeasibility of developing a universal protein analysis protocol. Furthermore, it must be considered that a metaproteome may include intracellular, extracellular and membrane-bound proteins, and ideally, the three protein fractions should be analysed for each sample. When possible, opting for gel-free protein fractionation seems to lead to a higher level of protein identification when compared with gel-based methods. For example, when analysing the metaproteome of activated sludge, the use of 2-DGE resulted in the identification of 38 proteins (Wilmes *et al.*, 2008a), while the 2D-nano-LC method led to the identification of 5029 proteins (Wilmes *et al.*, 2008b).

Table 1. Overview of selected metaproteomics studies

Environment	Number of peptides/proteins identified	Method	Databases	References
Human gut	NA/2214 proteins	LC-MS/MS	2 unmatched human gut metagenomes, several genomes from gut inhabitants and several nonhuman gut genome	Verberkmoes *et al.* (2009a)
Human gut	5010 peptides/NA	1D-PAGE, LC-MS/MS	Synthetic human gut metagenome (216 genomes from gut inhabitant) and 124 human gut unassembled nonannotated metagenomes	Rooijers *et al.* (2011)
Soil	NA/716 proteins	LC-MS/MS	Unmatched soil metagenome supplemented with 1606 genomes	Chourey *et al.* (2010)
Soil	NA/122	2D-PAGE, MALDI TOF/TOF MS/MS	Complete NCBInr, bacterial entries NCBInr and fungal entries NCBInr	Wang *et al.* (2011)
Marine	6533 peptides/1042 proteins	LC-MS/MS	SAR11 clade and specific microorganisms from Sargasso Sea metagenome as well as genomes from sequenced isolates	Sowell *et al.* (2009)
Marine	5389 peptides/2273 proteins	LC-MS/MS	Global Ocean Sampling combined metagenomes	Morris *et al.* (2010)
Freshwater	NA/1824 proteins	1D-PAGE, LC-MS/MS	Matched metagenomes	Lauro *et al.* (2011)
Acid mine drainage biofilm	NA/4107 proteins	LC-MS/MS	Biofilm_AMD_CoreDB database	Mueller *et al.* (2011)
Activated sludge	NA/5029 proteins	LC-MS/MS	Three distinct unmatched activated sludge metagenomes	Wilmes *et al.* (2008b)
Anaerobic digestion	NA/202 proteins	2D-PAGE, LC-ESI-MS/MS	Bacterial entries of the NCBI nonredundant database	Jehmlich *et al.* (2010)

1/2D-PAGE, one/two-dimensional polyacrylamide gel electrophoresis; LC-MS/MS, liquid chromatography-tandem mass spectrometry; MALDI-TOF, matrix-assisted laser desorption ionization-time of flight; ESI, electrospray ionization.

In addition, it is now apparent that metaproteomic approaches benefit from the availability of relevant metagenomic data, either matched or unmatched. As a result of this combined protocol, a new difficulty is encountered regarding the analysis and interpretation of the vast quantity of data generated. A major hurdle in the utilization of metagenomic data, impacting directly on metaproteomics, has been recognized as the assembly and the annotation of the collected genomic fragments. Rooijers *et al.* (2011) proposed an iterative workflow using nonannotated, unassembled metagenome sequences, which could be systematically used in metaproteomic investigations whenever relevant metagenomic data are available.

Future metaproteomic studies should aim to progress from proof of concept approaches to experimental designs leading to practical applications. For example, metaproteomic comparisons between healthy and diseased states within the human microbiome have yet to be carried out. Additionally, comparative *in situ* bioremediation investigations will need to be conducted to access the functional response of the natural mixed microbial communities to common pollutants. Furthermore, when investigating bioremediation processes, pollutants should be systematically monitored throughout the trials, to clarify the contaminant degradation status as a function of time. This in turn should allow for the proper assessment of the degrading abilities of the microbial communities investigated.

To conclude, the feasibility of metaproteomic studies has been successfully demonstrated in very diverse natural and engineered environments. However, only few studies to date employed this strategy to answer specific biological questions such as how complex communities define the biology of a given ecosystem (Ng *et al.*, 2010; Wilmes *et al.*, 2008b; Denef *et al.*, 2010; Mueller *et al.*, 2010; Lauro *et al.*, 2011), and it is now critical to move metaproteomics forward in order for this technology to achieve its full potential. To this end, future studies must be designed with an aim towards gaining some understanding of ecological concepts, and the data generated must be adequately analysed. In addition, attempts should be made to integrate such data in the context of system approaches to allow for the prediction of functional responses to environmental stimuli.

Acknowledgements

This work was conducted with the financial support of a University Fellow grant from the National University of Ireland, Galway.

References

Abram F, Enright AM, O'Reilly J, Botting CH, Collins G & O'Flaherty V (2011) A metaproteomic approach gives functional insights into anaerobic digestion. *J Appl Microbiol* **110**: 1550–1560.

Allen HK, Moe LA, Rodbumrer J, Gaarder A & Handelsman J (2009) Functional metagenomics reveals diverse β-lactamases in a remote Alaskan soil. *ISME J* **3**: 243–251.

Amann RI, Ludwig W & Schleifer KH (1995) Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol Rev* **59**: 143–169.

Awano S, Ansai T, Takata Y, Soh I, Akifusa S, Hamasaki T, Yoshida A, Sonoki K, Fujisawa K & Takehara T (2008) Oral health and mortality risk from pneumonia in the elderly. *J Dent Res* **87**: 334–339.

Bastida F, Moreno JL, Nicolás C, Hernández T & García C (2009) Soil metaproteomics: a review of an emerging environmental science. Significance, methodology and perspectives. *J Soil Sci* **60**: 845–859.

Bastida F, Nicolás C, Moreno JL, Hernández T & García C (2010) Tracing Changes in the microbial community of a hydrocarbon-polluted soil by culture-dependent proteomics. *Pedoshere* **20**: 479–485.

Bastida F, Jechalke S, Bombach P, Franchini AG, Seifert J, von Bergen M, Vogt C & Richnow HH (2011) Assimilation of benzene carbon through multiple trophic levels traced by different stable isotope probing methodologies. *FEMS Microbiol Ecol* **77**: 357–369.

Belnap CP, Pan C, VerBerkmoes NC, Power ME, Samatova NF, Carver RL, Hettich RL & Banfield JF (2010) Cultivation and quantitative proteomic analyses of acidophilic microbial communities. *ISME J* **4**: 520–530.

Benndorf D, Balcke GU, Harms H & Von Bergen M (2007) Functional metaproteome analysis of protein extracts from contaminated soil and groundwater. *ISME J* **1**: 224–234.

Benndorf D, Vogt C, Jehmlich N, Schmidt Y, Thomas H, Woffendin G, Shevchenko A, Richnow HH & von Bergen M (2009) Improving protein extraction and separation methods for investigating the metaproteome of anaerobic benzene communities within sediments. *Biodegradation* **20**: 737–750.

Bruneel O, Volant A, Gallien S *et al.* (2011) Characterization of the active bacterial community involved in natural attenuation processes in arsenic-rich creek sediments. *Microb Ecol* **61**: 793–810.

Burnum KE, Callister SJ, Nicora CD, Purvine SO, Hugenholtz P, Warnecke F, Scheffrahn RH, Smith RD & Lipton MS (2011) Proteome insights into the symbiotic relationship between a captive colony of *Nasutitermes corniger* and its hindgut microbiome. *ISME J* **5**: 161–164.

Chen S, Rillig MC & Wang W (2009) Improving soil protein extraction for metaproteome analysis and glomalin-related soil protein detection. *Proteomics* **9**: 4970–4973.

Chourey K, Jansson J, VerBerkmoes N, Shah M, Chavarria KL, Tom EL, Brodie EL & Hettich RL (2010) Direct cellular lysis/protein extraction protocol for soil metaproteomics. *J Proteome Res* **9**: 6615–6622.

Crielaard W, Zaura E, Schuller A, Huse S, Montijn R & Keijser B (2011) Exploring the oral microbiota of children at

various developmental stages of their dentition in the relation to their oral health. *BMC Med Genomics* **4**: 22.

Denef VJ, VerBerkmoes NC, Shah MB, Abraham P, Lefsrud M, Hettich RL & Banfield JF (2009) Proteomics-inferred genome typing (PIGT) demonstrates inter-population recombination as a strategy for environmental adaptation. *Environ Microbiol* **11**: 313–325.

Denef VJ, Kalnejais LH, Mueller RS, Wilmes P, Baker BJ, Thomas BC, VerBerkmoes NC, Hettich RL & Banfield JF (2010) Proteogenomic basis for ecological divergence of closely related bacteria in natural acidophilic microbial communities. *PNAS* **107**: 2383–2390.

Dewhirst FE, Chen T, Izard J, Paster BJ, Tanner ACR, Yu WH, Lakshmanan A & Wade WG (2010) The Human Oral Microbiome. *J Bacteriol* **192**: 5002–5017.

Donato JJ, Moe LA, Converse BJ, Smart KD, Berklein FC, McManus PS & Handelsman J (2010) Metagenomic analysis of apple orchard soil reveals antibiotic resistance genes encoding predicted bifunctional proteins. *Appl Environ Microbiol* **76**: 4396–4401.

Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, Samuel BS, Gordon JI, Relman DA, Fraser-Liggett CM & Nelson KE (2006) Metagenomic analysis of the human distal gut microbiome. *Science* **312**: 1355–1359.

Grant MM, Creese AJ, Barr G, Ling MR, Scott AE, Matthews JB, Griffiths HR, Cooper HJ & Chapple ILC (2010) Proteomic analysis of a noninvasive human model of acute inflammation and its resolution: the twenty-one day gingivitis model. *J Proteome Res* **9**: 4732–4744.

Guo T, Rudnick PA, Wang W, Lee CS, DeVoe DL & Balgley BM (2006) Characterization of the human salivary proteome by capillary isoelectricfocusing/nanoreversed-phase liquid chromatography coupled with ESI-Tandem MS. *J Proteome Res* **5**: 1469–1478.

Gygi SP, Rochon Y, Franza BR & Aebersold R (1999) Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol* **19**: 1720–1730.

Gygi SP, Rist B & Aebersold R (2000) Measuring gene expression by quantitative proteome analysis. *Curr Opin Biotechnol* **11**: 396–401.

Habicht KS, Miller M, Cox RP, Frigaard NU, Tonolla M, Peduzzi S, Falkenby LG & Andersen JS (2011) Comparative proteomics and activity of a green sulfur bacterium through the water column of Lake Cadagno, Switzerland. *Environ Microbiol* **13**: 203–215.

Hussain A, Al-Rawajfeh AE & Alsaraierh H (2010) Membrane bio reactors (MBR) in waste water treatment: a review of the recent patents. *Recent Pat Biotechnol* **4**: 65–80.

Janga SC, Díaz-Mejía JJ & Moreno-Hagelsieb G (2011) Network-based function prediction and interactomics: the case for metabolic enzymes. *Metab Eng* **13**: 1–10.

Jehmlich N, Schmidt F, von Bergen M, Richnow HH & Vogt C (2008) Protein-based stable isotope probing (Protein-SIP) reveals active species within anoxic mixed cultures. *ISME J* **2**: 1122–1133.

Jehmlich N, Kleinsteuber S, Vogt C, Benndorf D, Harms H, Schmidt F, von Bergen M & Seifert J (2010) Phylogenetic and proteomic analysis of an anaerobic toluene-degrading community. *J Appl Microbiol* **109**: 1937–1945.

Joshipura KJ, Hung HC, Rimm EB, Willett WC & Ascherio A (2003) Periodontal disease, tooth loss, and incidence of ischemic stroke. *Stroke* **34**: 47–52.

Kan J, Hanson T, Ginter J, Wang K & Chen F (2005) Metaproteomic analysis of Chesapeake Bay microbial communities. *Saline Systems* **1**: 7.

Keller M & Hettich R (2009) Environmental Proteomics: a paradigm shift in characterising microbial activities at the molecular level. *Microbiol Mol Biol Rev* **73**: 62–70.

Klaassens ES, de Vos WM & Vaughan EE (2007) Metaproteomics approach to study the functionality of the microbiota in the human infant gastrointestinal tract. *Appl Environ Microbiol* **73**: 1388–1392.

Lacerda CMR, Choe LH & Reardon KF (2007) Metaproteomic analysis of a bacterial community response to cadmium exposure. *J Proteome Res* **6**: 1145–1152.

Lauro FM, DeMaere MZ, Yau S *et al.* (2011) An integrative study of a meromictic lake Ecosystem in Antarctica. *ISME J* **5**: 879–895.

Lazarevic V, Whiteson K, Huse S, Hernandez D, Farinelli L, Østerås M, Schrenzel J & François P (2009) Metagenomic study of the oral microbiota by Illumina high throughput sequencing. *J Microbiol Methods* **79**: 266–271.

Lievens S, Eyckerman S, Lemmens I & Tavernier J (2010) Large-scale protein interactome mapping: strategies and opportunities. *Expert Rev Proteomics* **7**: 679–690.

Lo I, Denef VJ, VerBerkmoes NC *et al.* (2007) Strain-resolved community proteomics reveals recombining genomes of acidophilic bacteria. *Nature* **446**: 537–541.

Madsen EL (2011) Microorganisms and their roles in fundamental biogeochemical cycles. *Curr Opin Biotechnol* **22**: 456–464.

Mager DL, Haffajee AD, Devlinn PM, Norris CM, Posner MR & Goodson JM (2005) The salivary microbiota as a diagnostic indicator of oral cancer: a descriptive, nonrandomized study of cancer-free and oral squamous cell carcinoma subjects. *J Transl Med* **3**: 27.

Maron PA, Mougel C, Severine S, Houria A, Philippe L & Lionel R (2006) Protein extraction and fingerprinting optimisation of bacterial communities in natural environment. *Microbial Ecol* **53**: 426–434.

Mary I, Oliver A, Skipp P *et al.* (2010) Metaproteomics and metagenomic analyses of defined oceanic microbial populations using microwave cell fixation and flow cytometric sorting. *Microbial Ecol* **74**: 10–18.

Medina M & Sachs JL (2010) Symbiont genomics, our new tangled bank. *Genomics* **95**: 129–137.

Morris RM, Nunn BL, Frazar C, Goodlett DR, Ting YS & Rocap G (2010) Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *ISME J* **4**: 673–685.

Mueller RS, Denef VJ, Kalnejais LH *et al.* (2010) Ecological distribution and population physiology defined by proteomics in a natural microbial community. *Mol Syst Biol* **6**: 374.

Mueller RS, Dill BD, Pan C, Belnap CP, Thomas BC, VerBerkmoes NC, Hettich RL & Banfield JF (2011) Proteome changes in the initial bacterial colonist during ecological succession in an acid mine drainage biofilm community. *Environ Microbiol* **13**: 2279–2292.

Nelson KE, Weinstock GM, Highlander SK *et al.* (2010) A catalog of reference genomes from the human microbiome. *Science* **328**: 994–999.

Ng C, DeMaere MZ, Williams TJ *et al.* (2010) Metaproteogenomic analysis of a dominant green sulphur bacterium from Ace Lake, Antartica. *ISME J* **4**: 1002–1019.

Ong SE & Pandey A (2001) An evaluation of the use of two-dimensional gel electrophoresis in proteomics. *Biomol Eng* **18**: 195–205.

Pan C, Fischer CR, Hyatt D, Bowen BP, Hettich RL & Banfield JF (2011) Quantitative tracking of isotope flows in proteomes of microbial communities. *Mol Cell Proteomics* **10**: 1–11.

Park C & Helm RF (2008) Application of metaproteomic analysis for studying extracellular polymeric substances (EPS) in activated sludge flocs and their fate in sludge digestion. *Water Sci Technol* **57**: 2009–2015.

Park C & Novak JT (2007) Characterization of activated sludge exocellular polymers using several cation-associated extraction methods. *Water Res* **41**: 1679–1688.

Park C, Helm RF & Novak JT (2008a) Investigating the fate of activated sludge exocellular proteins in sludge digestion using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). *Water Environ Res* **80**: 2219–2227.

Park C, Novak JT, Helm RF, Ahn Y & Esen A (2008b) Evaluation of the extracellular proteins in full-scale activated sludges. *Water Res* **42**: 3879–3889.

Park JH, Lamb D, Paneerselvam P, Choppala G, Bolan N & Chung JW (2011) Role of organic amendments on enhanced bioremediation of heavy metal(loid) contaminated soils. *J Hazard Mater* **185**: 549–574.

Pradet-Balade B, Boulmé F, Beug H, Müllner EW & Garcia-Sanz JA (2001) Translation control: bridging the gap between genomics and proteomics? *Trends Biochem Sci* **26**: 225–229.

Qin J, Li R, Raes J *et al.* (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**: 59–65.

Raes J & Bork P (2008) Molecular eco-systems biology: towards an understanding of community function. *Nat Rev Microbiol* **6**: 693–699.

Ram RJ, VerBerkmoes NC, Thelen MP, Tyson GW, Baker BJ, Blake RC, Shah M, Hettich RL & Banfield JF (2005) Community proteomics of a natural microbial biofilm. *Science* **308**: 1915–1920.

Renella G, Chaudri AM & Brookes PC (2002) Fresh additions of heavy metals do not model long-term effects on microbial biomass and activity. *Soil Biol Biochem* **35**: 1203–1210.

Rodriguez-Brito B, Li L, Wegley L *et al.* (2010) Viral and microbial community dynamics in four aquatic environments. *ISME J* **4**: 739–751.

Röling WFM, Ferrer M & Golyshin PN (2010) Systems approaches to microbial communities and their functioning. *Curr Opin Biotechnol* **21**: 532–538.

Rooijers K, Kolmeder C, Juste C, Dore J, de Been M, Boeren S, Galan P, Beauvallet C, de Vos W & Schaap P (2011) An iterative workflow for mining the human intestinal metaproteome. *BMC Genomics* **12**: 6.

Rosselló-Mora R & Amann R (2001) The species concept for prokaryotes. *Microbiol Rev* **25**: 39–67.

Rudney JD, Xie H, Rhodus NL, Ondrey FG & Griffin TJ (2010) A metaproteomic analysis of the human salivary microbiota by three-dimensional peptide fractionation and tandem mass spectrometry. *Mol Oral Microbiol* **25**: 38–49.

Rusch DB, Halpern AL, Sutton G *et al.* (2007) The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol* **5**: 398–431.

Schulze WX, Gleixner G, Kaiser K, Guggenberger G, Mann M & Schulze ED (2005) A proteomic fingerprint of dissolved organic carbon and of soil particles. *Oecologia* **142**: 335–343.

Singleton I, Merrington G, Colvan S & Delahunty JS (2003) The potential of soil protein-based methods to indicate metal contamination. *Appl Soil Ecol* **23**: 25–32.

Solaiman Z, Marschner P, Wang D & Rengel Z (2007) Growth, P uptake and rhizosphere properties of wheat and canola genotypes in an alkaline soil with low P availability. *Biol Fertil Soils* **44**: 143–153.

Sowell SM, Wilhelm LJ, Norbeck AD, Lipton MS, Nicora CD, Barofsky DF, Carlson CA, Smith RD & Giovanonni SJ (2009) Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J* **3**: 93–105.

Sowell SM, Abraham PE, Shah M, Verberkmoes NC, Smith DP, Barofsky DF & Giovannoni SJ (2011) Environmental proteomics of microbial plankton in a highly productive coastal upwelling system. *ISME J* **5**: 856–865.

Toyoda A, Iio W, Mitsumori M & Minato H (2009) Isolation and identification of cellulose-binding proteins from sheep rumen contents. *Appl Environ Microbiol* **75**: 1667–1673.

Tringe SG, von Mering C, Kobayashi A *et al.* (2005) Comparative metagenomics of microbial communities. *Science* **308**: 554–557.

U.S. EPA. (2003) Environmental regulations and technology: control of pathogens and vector attraction in sewage sludge, EPA/625/R-92/013, revised July, 2003, Cincinnati, OH.

Venter CJ, Remington K, Heidelberg JF *et al.* (2004) Environmental shotgun sequencing of the Sargasso Sea. *Science* **304**: 66–74.

Verberkmoes NC, Russell AL, Shah M *et al.* (2009a) Shotgun metaproteomics of the human distal gut microbiota. *ISME J* **3**: 179–189.

Verberkmoes NC, Denef VJ, Hettich RL & Banfield JF (2009b) Systems biology: functional analysis of natural microbial consortia using community proteomics. *Nat Rev Microbiol* **7**: 196–205.

Vogel TM, Simonet P, Jansson JK, Hirsch PR, Tiedje JM, van Elsas JD, Bailey MJ, Nalin R & Philippot L (2009) TerraGenome: a consortium for the sequencing of a soil metagenome. *Nat Rev Microbiol* **7**: 252.

Wang HB, Zhang ZX, Li H *et al.* (2011) Characterization of metaproteomics in crop rhizospheric soil. *J Proteome Res* **10**: 932–940.

Williams MA, Taylor EB & Mula HP (2010) Metaproteomic characterization of a soil microbial community following carbon amendment. *Soil Biol Biochem* **42**: 1148–1156.

Wilmes P & Bond PL (2004) The application of two-dimensional polyacrylamide gel electrophoresis and downstream analyses to a mixed community of prokaryotic microorganisms. *Environ Microbiol* **6**: 911–920.

Wilmes P & Bond PL (2006) Metaproteomics: studying functional gene expression in microbial ecosystems. *Trends Microbiol* **14**: 92–97.

Wilmes P, Wexler M & Bond PL (2008a) Metaproteomics provides functional insight into activated sludge wastewater treatment. *PLoS ONE* **3**: e1778.

Wilmes P, Andersson AF, Lefsrud MG *et al.* (2008b) Community proteogenomics highlights microbial strain-variant protein expression within activated sludge performing enhanced biological phosphorus removal. *ISME J* **2**: 853–864.

Wilmes P, Bowen BP, Thomas BC, Mueller RS, Denef VJ, VerBerkmoes NC, Hettich RL, Northen TR & Banfield JF (2010) Metabolome-proteome differentiation coupled to microbial divergence. *mBio* **1**: e00246–10. doi:10.1128/mBio.00246-10.

Xie H, Onsongo G, Popko J, de Jong EP, Cao J, Carlis JV, Griffin RJ, Rhodus NL & Griffin TJ (2008) Proteomics analysis of cells in whole saliva from oral cancer patients via value-added three-dimensional peptide fractionation and tandem mass spectrometry. *Mol Cell Proteomics* **7**: 486–498.

Yooseph S, Sutton G, Rusch DB *et al.* (2007) The *Sorcerer II* global ocean sampling expedition: expanding the universe of protein families. *PLoS Biol* **5**: e16.

Zoetendal EG, Vaughan EE & de Vos WM (2006) A microbial world within us. *Mol Microbiol* **59**: 1639–1650.

Appendix References

DeLong, E.F. (1992) Archaea in coastal marine sediments. *Proc*. *Natl*. *Acad*. *Sci*. *USA* **89**, 5685-5689.

Lane, D., Pace, .B., Olsen, G.J., Stahl, d.A., Sogin, M.L. and Pace, N.R. (1985) Rapid determination of 16S ribosomal RNA sequences for phylogenetic analysis. *Proc*. *Natl*. *Acad*. *Sci*. *USA* **82**, 6955-6959.

Stackebrandt, E. and Goodfellow, M. (1991) Nucleic acid techniques in bacterial systematic. Wiley, England.

Yu, Y., Lee, C., Jaai, K. and Hwang, S. (2005) Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reactions. *Biotechnol*. *Bioeng*. **89**, 670-679.