How to use molecular biology tools for the study of the anaerobic digestion process?

Angela Cabezas, Juliana Calabria de Araujo, Cecilia Callejas, Amandine Gales, Jérôme Hamelin, Antonella Marone, Diana Z. Sousa, Eric Trably, Claudia Etchebehere

How to use molecular biology tools for the study of the anaerobic digestion process?

Angela Cabezas[1,2], Juliana Calabria de Araujo[3], Cecilia Callejas[1,4], Amandine Galès[6], Jérôme Hamelin[6], Antonella Marone[6], Diana Z. Sousa[5], Eric Trably[6], Claudia Etchebehere[1]

[1] Laboratorio de Ecología Microbiana, Departamento de Bioquímica y Genómica Microbiana, Instituto de Investigaciones Biológicas Clemente Estable, Av. Italia 3318, Montevideo, C.P. 11600, Uruguay.

[2] Laboratorio de Biotecnología, Facultad de Ingeniería, Universidad ORT Uruguay, Cuareim 1451, Montevideo, C.P. 11100, Uruguay.

[3]Departamento de Engenharia Sanitária e Ambiental, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Brazil.

[4]Laboratorio de Bioprocesos Ambientales, Institutito de Ingeniería Química. Facultad de Ingeniería, Universidad de la República, Herrera y Reissig 565, Montevideo, Uruguay.

[5]Laboratory of Microbiology, Wageningen University, Dreijenplein 10, 6703 HB Wageningen, The Netherlands.

[6]Laboratoire de Biotechnologie de l'Environnement, INRA, UR0050, Avenue des Etangs, F-11100, Narbonne, France.

Corresponding author:

Claudia Etchebehere

e-mail: cetchebehere@iibce.edu.uy.

Telephone: (598) 2487-1616. Fax: (598) 2487-5461.

Abstract

Anaerobic Digestion is used with success for the treatment of solid waste, urban and industrial effluents with a concomitant energy production. The process is robust and stable, but the complexity of the microbial community involved in the process is not yet fully comprehensive. Nowadays, the study of this complex ecosystem is facilitated by the availability of different molecular tools, but it is very important to choose the adequate tool to answer specific questions. The aim of this review is to describe different molecular techniques, indicate the

questions that can be addressed by each technique, enumerate their limitations and give practical advices for their use. Examples of how the molecular tools have been used to address various questions in anaerobic digestion are presented. The key point now is to apply all this information to improve anaerobic digestion. The integration of concepts of microbial-ecology, environmental-engineering, modeling and bioinformatics is currently necessary.

Key words

Anaerobic digestion, bioreactor, molecular biology, microbial ecology, function, microbial resource management.

List of abbreviations

AD	anaerobic digestion
ANOSIM	analysis of similarity
ANOVA	analysis of variance
ARDRA	amplified ribosomal DNA restriction analysis
BES	2-bromoethanesulphonic acid
BONCAT	biorthogonal non-canonical amino acid tagging
CE-SSCP	capillary electrophoresis Single Strand Conformation Polymorphism
CCA	canonical correspondence analysis
CLSM	confocal laser scanning microscopy
CNV	copy number variation
DIET	direct interspecies electron transfer
DGGE	denaturing gradient gel electrophoresis
DNA	desoxyribo nucleic acid
EGSB	Expanded granular sludge blanket
FISH	fluorescent in situ hybridization
HIT	Hydrogen interspecies transfer
MAR-FISH	micro auto radiographic fluorescent in situ hybridization
NanoSIMS	nanoscale secondary ion mass spectrometry

NGS	next generation sequencing
NMDS	nonmetric multidimentional scaling
NPMANOVA	non parametric multivariate ANOVA
OLR	organic loading rate
OTU	operational taxonomic unit
PCA	principal component analysis
PCoA	principal coordinate analysis
PCR	polymerase chain reaction
qPCR	quantitative polymerase chain reaction
RISA	ribosomal intergenic spacer analysis
RNA	ribo nucleic acid
rRNA	ribosomal RNA
SAO	syntrophic acetate oxidizers
SIMSISH	secondary ion mass spectrometry combined with FISH
SIP	stable isotope probing
SSCP	Single Strand Conformation Polymorphism
ssDNA	single strand DNA
T-RFLP	Terminal-Restriction Fragment Length Polymorphism
T-RF	Terminal-Restriction Fragment
UASB	Upflow anaerobic sludge blanket
VSS	volatile suspended solids

1-Introduction

1.1 Anaerobic Digestion, a robust and complex process

Anaerobic Digestion (AD) of waste and effluents is a robust process that is nowadays used with success in full scale systems for the treatment of solid waste and urban and industrial wastewaters. Worldwide, there are thousands of high rate sludge bed reactors for industrial

wastewater treatment, and millions of domestic biogas plants generating energy from waste (Ren 2013; Noyola et al. 2012). Many engineering enterprises are involved in the design and installation of these type of biological reactors. Moreover, new technologies are being tested at lab-scale using different reactor configurations (e.g. UASB, EGSB, anaerobic membrane reactors), integrating different treatments (N-removal, S-removal, micropollutants removal, etc.), or systems to recover nutrients (Batstone and Virdis 2014).

Anaerobic biological process is driven by a complex network of microorganisms belonging to *Bacteria* and *Archaea* domains, working together for a complete degradation of the organic compounds into CH_4 and CO_2. In the complete anaerobic chain, four main metabolic steps are involved: 1-hydrolysis, 2-fermentation, 3- acetogenesis, 4- methanogenesis (Zinder 1984) (Figure 1). These four steps involve different microbial guilds who are specialized in different processes. A diverse number of hydrolytic and fermentative bacteria take part in the first two steps, whereas the oxidation of intermediate fermentation products to acetate is performed by either hydrogen- or formate- producing acetogens (Stams and Plugge 2009). Ultimately, methane is exclusively generated from acetate and hydrogen/CO_2 by methanogenic archaea.

In general, the microbial composition of anaerobic digestion sludge is highly diverse and shows high redundancy. This means that several microorganisms are metabolically flexible and capable of performing the same work. It has been postulated that this characteristic is one of the reasons for the robustness of anaerobic digestion processes (Fernández et al. 1999; Zumstein et al. 2000).

To study these very complex ecosystems there are nowadays different molecular tools available. It is very important to choose the adequate tool to answer the questions formulated in each experiment.

1.2 Questions to answer

Regarding the microbial composition and function of the communities in methanogenic bioreactors, four main questions are usually formulated:

1-Who is there?, 2- How does the community change over time?, 3-How many microorganisms from the different groups are present?, 4-What are the specific functions of the microorganisms within the community and their relation to each other?

Molecular ecology techniques have been evolving over the past years to give answers to these questions and an array of options are available and will be discussed in this review (Figure 2).

1.2.1 Microbial diversity - Who is there?

Information on the diversity and identity of the microorganisms involved in the anaerobic digestion process is important to understand bioreactor functioning, especially when concerning new metabolic processes. The discovery of microorganisms involved in the anaerobic oxidation of ammonium (Anammox process) is a nice example (Jetten et al. 1999; Ni and Zhang 2013). Other important key process are syntrophic oxidation of organic acids (McInerney et al. 2008), degradation of recalcitrant compounds such as detergents (Okada et al. 2014), and anaerobic oxidation of methane (Fernández et al. 2008).

To assign the identity of the microorganisms involved in a microbial community the most frequently used technique is based on the analysis of the 16S rRNA gene. This gene was proposed as a housekeeping genetic marker to study bacterial and archaeal phylogeny and taxonomy for several reasons: (i) is present in almost all bacteria and archaea; (ii) the function over time has not changed, suggesting that random sequence changes are a more accurate measure of time (evolution); and (iii) the 16S rRNA gene (1,500 bp) is large enough for informatics purposes (Patel et al. 2001). Nowadays, it is possible to determine the genus and species of a bacteria or archaea by sequencing the 16S rRNA gene and compare the sequence with available databases. This technique was adapted to study the microbial composition of a sample as will be explained below.

1.2.2 Microbial dynamics - How does the community change over time?

Different reactor operation parameters, such as applied organic loading rate, hydraulic retention time or operating temperature, are frequently tested to determine the optimal operation conditions for the anaerobic digestion process. In order to explain how these operation conditions can change reactor performances it is necessary to monitor the microbial community composition during operation. When the objective of the research is to test different reactor configurations, different sources of inocula, or different substrates, comparing the microbiology of different systems is needed. In all these cases, numerous community structures could be obtained and the application of a fingerprinting technique such as DGGE, T-RFLP or SSCP is a suitable choice.

1.2.3 Microbial quantification -How many microorganisms from different groups are present?

In AD process it is important to quantify some key groups, in particular the density and proportion of methanogens because of their relevance for ensuring an efficient methanogenic

process. For this, techniques that quantify different groups of microorganisms present in a complex community such as Fluorescent *in situ* Hybridization (FISH) or quantitative PCR (q-PCR) are adequate.

1.2.4 Microbial function –What are the roles of different microbial groups in anaerobic microbial communities?

Discovering specific microbial roles in anaerobic microbial communities is currently one of the most challenging issues for microbiologists and molecular ecologists. Metabolic pathways and interspecies relations involved in the anaerobic process are frequently unknown, in particular when a difficult process is going on, such as the degradation of recalcitrant compounds. Knowing the identity of the microorganisms involved in the process (section 1.2.1) can give a hint on the metabolic potential of abundant players, but it is generally not enough to assign a function to those microorganisms. It is also possible that a single microorganism plays a role at different steps of the metabolic pathways, e.g. *Clostridium* sp. can encode both hydrolytic and fermentative enzymatic machineries. In these cases more sophisticated techniques, such as proteomics, metagenomics, metabolomics and techniques that link the identity with function, such as Stable Isotope Probing (SIP) or micro-autoradiography linked to FISH (MAR-FISH), are the most relevant choice.

2-Sampling, storage of samples and environmental data collection

2.1 Sampling

Sampling is one of the most important steps in microbial ecology analysis. A good sampling strategy is necessary to ensure the success of the study. It has to be taken into account that the use of statistical tools to compare microbial communities might not be correctly applied if a representative sampling is not performed from the beginning of the study.

Choice of sample amount is not trivial and has to be considered from the beginning of the study, both to ensure enough biomass concentration for molecular analysis and a representative selection of the microorganisms in the reactor sludge. The majority of the techniques and kits used to extract DNA or RNA from reactor's biomass are designed for soil samples (Griffiths et al. 2000), e.g. PowerSoil DNA isolation (Mo Bio Laboratories Inc. Carlsbad, CA) and Nucleo-Spin Soil (Macherey-Nagel, Düren, Germany). Considering these example kits, manufacturers protocols suggest the use of 0.5 g to 1.0 g of soil. Reactor's sludges are in general suspended in the liquid mixture and present fewer amounts of inorganic compounds and higher amount of cells than soil. According to that, the volume of the sample has to be recalculated. The amount

and frequency of sampling might also have an impact in the biological reactor itself; volume of sampling needs to be considered to avoid volume changes in the reactor, especially when working with small lab-scale reactors. To guarantee bioreactor sample representativeness, it is recommended to take samples from different parts of the reactor and pool them, especially in the case of full scale reactors. The sampling of the biomass from solid waste digesters may present additional difficulties due to the heterogeneity of the system. If statistical analysis is going to be applied it is recommended to plan the sampling with a specialist. When the microbial community is analyzed in time series, it is important to adapt the sampling frequency with the objective that samples represent the studied period. Several samples taken during a phase of operation are recommended to be able to compare the different phases. Last but not least, sampling directly in the vial that will be used for further analysis spares some manipulations and limits the risk of sample loss.

Once the sample is taken it must be stored correctly to avoid microbial growth during storage. The recommended temperature of storage is -20ºC for further DNA extraction and -80ºC if it is planned to work with RNA. If a freezer is not available close to the place of sampling, it is recommended to store the samples on ice until you reach the laboratory. If the samples are going to be analyzed by FISH, they should be fixed with formaldehyde or ethanol before storing at -20ºC.

A fact to consider before DNA/RNA extraction is how to handle the sample. Depending on the kind of reactor, the biomass can be suspended, attached to an inert support, aggregated in granules or forming flocs. In some cases, before nucleic acids extraction, it will be necessary to detach the biomass from the support or disrupt the aggregates. Several protocols are available in literature using ultrasound treatment (Perna et al. 2011) or physical disruption (Bergmann et al. 2010).

2.2 Environmental data collection

In order to be able to link reactor performances with the microbial community structure, collecting reactor operation data at sampling time is absolutely necessary. In general, parameters such as temperature and pH are recorded daily, but other parameters that need further physicochemical analysis might be determined weekly or monthly depending on the duration of reactor operation. Information about reactor design and wastewater composition is also very important for interpreting the results. There is no general rule but the collection of a complete set of operating data is preferred.

3- Tools for studying microbial diversity (Who is there?)

3.1 Cloning and Sanger sequencing

The identity of the microorganisms present in a sample can be determined through PCR amplification and analysis of conserved marker genes. The 16S rRNA gene is the most widely used marker gene and has the most extensive reference database (Su et al. 2012). During the last decades cloning in a plasmid vector followed by Sanger sequencing has been widely used. The technique involves PCR amplification of the 16S rRNA genes using primers directed to either *Bacteria* or *Archaea* domains. In the anaerobic digestion research area, *Bacteria* and *Archaea* are both important, and both corresponding clone libraries can be made separately by choosing domain-specific PCR primer sets during the initial PCR amplification step (Vanwonterghem et al. 2014).

Several *Bacteria* and *Archaea* specific primers have been designed (Lane 1991; Grosskopf et al. 1998; Leclerc et al. 2004), e.g. the *Bacteria* specific PCR primer sets 27f and 1492r (Lane 1991) and the *Archaea*-specific primers A349f (Takai and Horikoshi 2000) and A915r (Stahl and Amann 1991).

The sequences obtained are then clustered into arbitrary Operational Taxonomic Units (OTUs), generally defined as sequences that shared more than 97% identity. This value has been chosen considering that, in general, two strains from the same species share more than 97% homology in their 16S rRNA gene sequences (Yarza et al. 2014). The taxonomic position of the representative sequences is determined by comparing the sequences with databases using BLAST at National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/) (Altschul et al. 1990), the Classifier at the Ribosomal Database Project (RDP) (https://rdp.cme.msu.edu/) (Wang et al. 2007) or the European database (http://www.ebi.ac.uk/ena).

Several ecological questions of the AD process have been answered by using 16S rRNA gene sequencing techniques. At the beginning, as the cost by sequence was high, the clones were previously analyzed by other techniques (Amplified Ribosomal DNA Restriction Analysis (ARDRA), Single Strand Conformation Polymorphism (SSCP)) and only representative clones were then sequenced. Studies were focused on describing the composition of the complex communities developed in lab-scale methanogenic reactors (Godon et al. 1997) or full scale systems (Roest et al. 2005). Other works focused on the stability of these systems (Fernández et al. 1999, Zumstein et al. 2000) or comparing communities developed under different temperatures (Sekiguchi et al. 1998) or feeding conditions (Goberna et al. 2009).

With the development of the automatic Sanger sequencing technology, higher number of clones could be sequenced. In a very extensive sequencing effort, Riviére et al. (2009) analyzed the

microbial composition of seven anaerobic sludge digesters. A total of 9,890 16S rRNA gene sequences were analyzed. From the comparison of all the retrieved information, the authors defined three component models within the bacterial communities: 1-a core group of phylotypes common to most of the digesters, 2-phylotypes shared among few digesters and 3-specific phylotypes (Table 1).

To give a global view of the microbial diversity involved in the AD process in general, Nelson et al. (2011), compiled the information produced by the cloning/sequencing approach. The authors performed a meta-analysis based on all publicly available 16S rRNA gene sequences generated by Sanger sequencing from various anaerobic digesters up to may 2010. A total of 19,388 sequences (16,519 bacterial and 2,869 archaeal) were analyzed, representing 28 known bacterial phyla (e.g. *Proteobacteria, Firmicutes, Bacteroidetes, Chloroflexi)* (Table 1). Archaeal sequences were assigned to 296 OTUs, primarily *Methanosaeta* and the uncharacterized WSA2 group.

The cloning/sequencing approach is also very valuable to get sequence information for the design of signature oligonucleotides that are complementary to target groups, at the taxonomic level of family, genus, species, or strain. These newly designed oligonucleotides may serve in FISH studies and in the development of real-time PCR assays for quantification (Ariesyady et al. 2007).

The methodology might also be applied to study specific functional genes. As an example, one of the key genes involved in the methanogenic pathway was used to study the specific composition of this population (Zhu et al. 2011). A different strategy could be applied to study a specific population using 16S rRNA gene primers specifically designed to target that microbial group (Chouari et al. 2005).

3.2 Next generation sequencing of PCR amplicons

The development of Next Generation Sequencing (NGS) technologies has overcome three of the main limitations of the cloning/sequencing technology, as these new methods are: faster, cheaper and high-throughput. Millions of sequences can be obtained in a single run in a complete automatic method within a few hours or days. Sequencing technologies are based on the detection of incorporated nucleotides by different chemistries. In all cases, sequences are generated without the need of a conventional, vector-based cloning procedure (Shokralla et al. 2012). A complementary approach is the adoption of nucleotide barcodes in the amplification primers for multiplexing samples. In this way, up to 120 samples from different origins can be mixed in one run and after sequencing their data can be separated according to their barcode (Parameswaran et al. 2007; Meyer et al. 2007). This approach decreases the cost per sample

since more samples can be pooled in the sequencing run rather than sequencing fewer samples to greater depth (Cardenas and Tiedje 2008).

The first step is the PCR of the desired gene, as the length of the sequences obtained is shorter than in the Sanger methods, specific primers for this technology are used (Cardenas and Tiedje 2008, Wang and Qiang 2009). This step is followed by high-throughput sequencing of the resulting amplicons libraries by means of a selected NGS platform. There are four commercially available platforms which use PCR based sequencing systems: Roche 454 Pyrosequencing Genome Sequencer (Roche Diagnostics Corp. Branford, CT, USA), MiSeq and HiSeq 2000 (Illumina Inc. San Diego, CA, USA), AB SOLiD System (Life Technologies Corp. Carlsbad, CA, USA) and Ion Personal Genome Machine (Life Technologies, South San Francisco, CA, USA) (Shokralla et al. 2012).

These new sequencing technologies are based on different concepts: 454 Pyrosequencing, Ion Personal Genomic Machine (Ion PCM) and Illumina uses real-time sequencing-by-synthesis. In these technologies each nucleotide incorporated by DNA polymerase generates a signal detected by a specific detector system. In the case of 454 Pyrosequencing each nucleotide incorporated release a pyrophosphate molecule which is linked to the production of light by the action of the enzyme luciferase (Margulies et al. 2005). The Ion Personal Genome Machine detects the changes in the hydrogen ion concentration produced when a nucleotide is incorporated into a strand of DNA by the polymerase action (Rothberg et al. 2011). The Illumina platform sequencing is based on the detection of fluorescent signal release after incorporation of a modified nucleotide (Shendure and Ji 2008). These three techniques includes the immobilization of the library fragments on beads (454 and Ion PGM) or a surface flow cell (Illumina) whose surfaces carry oligonucleotides complementary to specific adapter sequences ligated or PCR-generated onto both ends of the fragmented library. After a step of amplification, the plate/chip/flow cell is sequenced en masse by the instrument.

Unlike the previous three platforms, SOLiD uses a sequencing by-oligo ligation technology. This process couples oligo adaptor-linked DNA fragments with complementary oligos immobilized on the surface of 1-mm magnetic beads. After the ligation step, a fluorescent readout identifies the ligated 8-mer oligo, which corresponds to one of the four possible bases (Shokralla et al. 2012).

Each technology has advantages and disadvantages, the major advantages of the 454 Pyrosequencing platform are the long read length obtained and its relatively short run time. Longer sequences generated through 454 provide higher flexibility in terms of accurate annotation of reads in ecological applications involving non model organisms. This has made 454 the most commonly used NGS platform for the analysis of environmental DNA for

ecological applications. The Ion PGM present a cheap alternative with relative long reads (>200 bp) but lower coverage than 454-pyrosequencing. The advantage of both Illumina and SOLiD systems is the high output per run compared to 454 pyrosequencing and Ion PGM. The main drawback of these systems is the relative short-read length. Then, before choosing a platform, it is necessary to evaluate the need of higher coverage or higher sequences length (Shokralla et al. 2012).

The interpretation of the huge amount of produced data requires appropriate bioinformatic tools and a specific know-how in addition to efficient computational resources. Three bioinformatic pipelines are widely used to analyze high amounts of sequences: QIIME (Caporaso et al. 2010), Mothur (Schloss et al. 2009) and the online tool developed by the Ribosomal Database Project (http://pyro.cme.msu.edu/) (Cole et al. 2014).

Although this technology was developed ten years ago (Margulies et al. 2005), several applications to the analysis of microbial communities from anaerobic digesters have been reported over the past years. As the sequencing cost has decreased markedly the technologies are more popular and several samples can be analyzed in a short time with reasonable costs. This has allowed to perform studies comparing the microbiology of high number of reactors operated during long time periods. Werner et al. (2011), analyzed the microbiology of 9 full scale methanogenic bioreactors treating brewery wastewater by 454-pyrosequencing. The high amount of datasets was analyzed using specialized bioinformatic tools and correlated with environmental data using statistical analysis. The authors found that each of the nine facilities had a unique community structure with an unprecedented level of stability. Syntrophic populations were highly stable, resilient, and specific for function and environment. This indicates that resilience, rather than dynamic competition, plays an important role in maintaining the necessary syntrophic population. They also found that communities with a greater phylogenetic variability functioned more efficiently.

Similarly, Sundberg et al. (2012) studied the bacterial and archaeal compositions of 21 full scale biodigesters operated in thermophilic or mesophilic conditions to produce energy using different solid waste. The reactors were fed with sewage sludge and various combinations of sewage sludge in codigestion with other wastes. The predominant populations in sewage sludge-fed reactors included *Actinobacteria*, *Proteobacteria*, *Chloroflexi*, *Spirochetes*, and *Euryarchaeota*, while *Firmicutes* was the most prevalent in the codigestion reactors fed with other waste (Table 1). The main bacterial class found in all reactors was *Clostridia*. Acetoclastic methanogens were detected in sewage sludge, and not in other reactors. Their absence suggests that methane formation from acetate takes place mainly via syntrophic acetate oxidation in these reactors, as previously reported (Karakashev et al. 2006). Statistical analysis revealed that the microbial composition was mainly governed by the type of substrate and the process temperature.

Other more specific projects studied the relationship of the microbial communities with different environmental variables as fed sources (Ziganshin et al. 2013) or temperature (Tian et al. 2015).

The principal advantage of the sequencing-based methods is that the identity of the microorganisms present in a sample can be determined. The Sanger sequencing method gives more accurate taxonomic assignments as full length sequences can be retrieved (1,500 bp), but, the method is onerous, time-consuming and nowadays widely replaced by NGS methods. However, the recent improvements of the NGS equipment allow achieving longer sequences maintaining a low cost. For instance, in the new 454 GS FLX+System (Roche) reads up to 1,000 bp can be obtained. The bottleneck of the new methodology is now to develop bioinformatic tools to analyze the high amount of data generated. The tools available require informatic skills and important computer capacity.

Functional insight into the process can be inferred from 16S-based studies by searching for closely related cultured species. However, this should be done with caution as closely related organisms can be functionally different (Jaspers and Overmann 2004). Metabolic functionality can also be inferred by correlating community composition and dynamics with operational conditions and performance parameters using multivariate statistical tools. This correlation is particularly important in anaerobic digestion, where niche differentiation occurs and syntrophic groups are responsible for fulfilling essential steps in the process (Vanwonterghem et al. 2014).

3.3 DNA microarrays

DNA microarrays or microchips, are a fast, semi-quantitative technique for phylogenetic identification of bacterial and archaeal species or for detection of functional genes. It is based on the hybridization between extracted DNA sample or 16S rRNA amplicons, which are fluorescently labeled, and complementary oligonucleotide probes that are immobilized on a glass slide (microchip). When hybridization occurs, fluorescence can be detected using a laser. DNA microarrays can detect thousands of genes at the same time (high-throughput) and enables screening of microbial structure and activities (Čater et al. 2013). Morover, it is fast and once the chip has been design and fabricated the cost of applying the technique is not high. The main disadvantage is that the genes detected are limited to genes previously included in the microchip. Taking into account this limitation, the technique is very useful to compare communities from samples taken under different conditions.

To study AD, an ANAEROCHIP was used targetting methanogens which includes 103 probes. This chip has been used to investigate the methanogenic community from a thermophilic

continuously stirred tank reactor (Franke-Whittle et al. 2009) and to follow the start up of solid waste digesters (Goberna et al. 2015).

Functional genes can be studied using a special chip which particularly includes different genes from the environment (as the GeoChip) (Tu et al. 2014). This strategy was used to study changes in the microbial community during operation of a combined nitritation-anammox reactor established to treat anaerobic digestion supernatant (Zhao et al. 2014). As far as we know, there is no available microarray especially designed with functional genes present in anaerobic digesters communities. This functional microarray may have high potential application to monitor bioreactors operation and start up.

4. Microbial dynamics (How does the community change over time?)

4.1 Fingerprinting techniques: DGGE, T-RFLP, SSCP

The most common fingerprinting methods include Denaturing Gradient Gel Electrophoresis (DGGE), Single-Strand Conformation Polymorphism (SSCP), Terminal-Restriction Fragment Length Polymorphism (T-RFLP) and Ribosomal Intergenic Spacer Analysis (RISA). These methods are based on the analysis of PCR products amplified from 16S rRNA genes (DGGE, SSCP and T-RFLP) or from the ribosomal intergenic region between 16S and 23S rRNA genes (RISA). With these methods, a community fingerprint based on sequence polymorphism is generated which gives an assessment of community structure and fluctuation over time in ecological studies (Schmalenberger and Tebbe 2002; Rastogi and Sani 2011). Fingerprints of the active part of a community may also be obtained after RNA extraction and reverse transcription of the 16S rRNA (Delbès et al. 2000; Delbès et al. 2001). The possibility to get fingerprinting patterns from both 16S rDNA and 16S rRNA confers a great advantage to DGGE, SSCP and T-RFLP with respect to the RISA method. Other genes can be targeted with fingerprinting methods and it is not restricted to ribosomal genes (Figure 3). In anaerobic digestion, monitoring the activity of methanogens is a primary target. The *mcrA* gene carried by methanogens, which encodes the key enzyme called Methyl Coenzyme M Reductase catalysing the last step of methane formation, was established as a molecular marker (Lueders et al. 2001) since all known methanogens express the Mcr enzyme (Ferry, 1999). Nowadays, a satisfactory number of databank entries for *mcrA* gene are available, and sequence analyses reveal a huge diversity of methanogens. The presence of this enzyme provides a reliable diagnostic of methanogenesis in diverse environments (Reeve et al. 1997; Luton et al. 2002; Steinberg and Regan 2009).

437

438 *4.1.1 DGGE as one of the most popular and used molecular tool for monitoring microbial*
439 *community structure and dynamics in bioreactors*

440 DGGE is one of the most popular and used techniques for biodiversity assessment in bioreactor
441 samples (Boon et al. 2002; Arooj et al. 2007; Connaughton et al. 2006; Miura et al. 2007;
442 Pereira et al. 2014; Ramos et al. 2014).

443 DGGE protocols are relatively simple and straightforward and produce results in relatively short
444 time: after extraction of genomic DNA or RNA, target gene sequences (mostly 16S rRNA gene
445 fragments, or functional genes) are amplified using specially designed primers with GC clamps
446 in the PCR reaction, and PCR amplicons of equal length are separated electrophoretically in a
447 denaturing gradient gel (Muyzer et al. 1993) (Figure 3). The interpretation of DGGE gels is
448 rapid and straightforward, if only band patterns and relative band intensities are of interest. This
449 renders this method popular for fast comparative analysis of communities from different
450 reactors (Boon et al. 2002; Ramos et al. 2014) or from the same reactor operated under different
451 conditions (Arooj et al. 2007; Miura et al. 2007; Luo and Angelidaki 2012; Ramos et al. 2013;
452 Shi et al. 2013; Pereira et al. 2012). Sequencing of the DNA bands excised from gels allows the
453 identification of different members of the microbial community and further phylogenetic
454 analysis (Arooj et al. 2007; Connaughton et al. 2006; Díaz et al. 2011; Luo and Angelidaki
455 2012; Ramos et al. 2013; Ramos et al. 2014; Pereira et al. 2012).

456 Nevertheless, DGGE as any other techniques has some disadvantages and limitations. Reports
457 on methodological difficulties of DGGE include issues such as comigration of DNA from
458 different species in the same band (Sekiguchi et al. 2001b; Speksnijder et al. 2001) and
459 formation of multiple bands in the amplification of genes from single genomes (Brosius et al.
460 1981; Nübel et al. 1996). Moreover, this method has limitations with respect to its resolving
461 power, as typically only up to 50 bands can be distinguished in a gel lane and it is clearly
462 limited to the dominant members of the microbial community (with a threshold of roughly 0.1%
463 of the total) (Van Elsas and Boersma 2011). In spite of these reports, many researchers employ
464 this technique for the identification of important community members and dominant organisms
465 found in bioreactors.

466 The number of bands is assumed to accurately reflect the diversity of microbes in the sample
467 (Miura et al. 2007; Luo and Angelidaki 2012; Ziembinska-Buczynska et al. 2014) whilst the
468 relative intensity of each band is thought to reflect the relative abundance of the particular
469 organism represented by the band in the community (Dar et al. 2005; Connaughton et al. 2006;
470 Díaz et al. 2006; Shi et al. 2013).

Since identification of important community members as well as numerically most dominant members of a community is a key aspect of microbial community analysis of bioreactor samples, it is very important to know whether DGGE is a reliable technique to obtain such community data. In a systematic study conducted by Araujo and Schneider (2008) with artificial consortia, this technique was tested under conditions where results would not be affected by differences in DNA extraction efficiency from cells. They demonstrated that DGGE was suitable for identification of all important community members in the three-membered artificial consortium, but not for identification of the dominant organisms in this small community. Multiple DGGE bands obtained for single organisms with the V3 primer pair (targeting 16S rRNA) could greatly confound interpretation of DGGE profiles.

Despite all the limitations mentioned above, PCR-DGGE has turned into a routine fingerprinting method to study microbial diversity because it has a reasonable resolving power and allows fast comparative analysis between different samples (Van Elsas and Boersma 2011).

DGGE has been used to analyze the biodiversity of granular sludge (Roest et al. 2005; Zhang et al. 2005), to determine the microbial composition and structure of different types of granules in a UASB reactor from an industrial brewery (Díaz et al. 2006) and in a UASB reactor treating domestic wastewater (Pereira et al. 2012), to compare the start-up and evolution of mesophilic and thermophilic UASB reactors (LaPara et al. 2000; Liu et al. 2002), to investigate changes in the microbial population in continuously stirred tank reactors (Rincon et al. 2006; Ueno et al. 2001), and to analyze the microbial population dynamics in anaerobic reactors treating landfill leachate or trichloroethane (Calli et al. 2006; Tresse et al. 2005). DGGE was also used to investigate the effect of operating temperatures on the microbial community profiles in a high cell density hybrid anaerobic reactor (Kundu et al. 2012), to analyze the effect of the carbon source on the microbial community structure and performance of an anaerobic reactor (Kundu et al. 2013a), and to evaluate changes in microbial communities in a hybrid anaerobic reactor with organic loading rate and temperature (Kundu et al. 2013b).

The correlation of microbial community structure with wastewater composition and reactor's performances was investigated by Kundu et al. (2013a). Self-immobilized granules were developed in synthetic wastewaters based on different carbon sources (glucose, sugarcane molasses, and milk), in three hybrid anaerobic reactors operated at 37°C. To study archaeal community structure, a polyphasic approach was used with both qualitative and quantitative analysis. While DGGE of 16S rRNA gene did not reveal major shifts in diversity of *Archaea* with change in substrate, quantification of different groups of methanogens and total bacteria by real-time PCR showed variations in relative abundances with the dominance of *Methanosaetaceae* and *Methanobacteriales*. These data were supported by differences in the ratio of total counts of *Archaea* and *Bacteria* analyzed by catalyzed reporter deposition –

fluorescence in situ hybridization. During hydraulic and organic shocks, the molasses-based reactor showed the best performances followed by the milk-and glucose-based reactors. This study indicates that carbon source shapes the microbial community structure more in terms of relative abundance with distinct metabolic capacities rather than its diversity.

Ramos et al. (2014) used DGGE to evaluate the impact of O_2 on microbial communities in an industrial-pilot scale sewage sludge reactor. Ziembinska-Buczynska et al. (2014) analyzed the bacterial community performing methanogenesis in a pilot scale anaerobic chamber during the shift from mesophilic to thermophilic conditions to point at the group of temperature tolerant microorganisms and their performances, by using DGGE. Taken together, these examples indicate that DGGE has turned into a routine fingerprinting method and is still one of the most used molecular methods to study microbial diversity in anaerobic reactors.

4.1.2 T-RFLP a molecular tool to monitor bacterial and archaeal communities in anaerobic reactors

T-RFLP emerged as a molecular tool for the study of microbial communities when Liu (1997) adapted an existing method (RFLP) in order to perform a rapid analysis of complex environments. T-RFLP analysis involves amplification of target genes from whole-community DNA extracts by using specific primer pairs, one of which is fluorescently labeled. Subsequently, amplicons are digested with restriction enzymes, typically using 4-base cutters, and fragments are size separated and detected on automated sequencers (Figure 3). As only terminal fragments (T-RFs) are detected the complexity of the profiles is considerably reduced. The polymorphism depends on the fragment length and an internal size standard labeled with a different fluorophore running with each sample allows a precise assignment of the fragment lengths with single base pair resolution. The patterns of T-RFLP for a sample can be converted into a numerical form and allows, after applying standardization procedures, to compare samples using a variety of statistical approaches. In this way, by using T-RFLP profiles, questions related to the diversity and dynamics of complex ecosystems might be revealed. Moreover, a presumptively phylogenetic assignment of T-RFs might be possible relating the T-RFLP profiles to *in silico* restriction of OTUs obtained from cloning libraries of the same or related samples. This is especially useful in simple ecosystems or when a specific population is targeted, where the number of *in silico* T-RFs retrieved from the OTUs of a cloning library is low. For more complex communities, the taxonomic assignment for all T-RF might not be possible. Nonetheless, the predominant T-RFs are generally identified. One shortcoming is that discrepancies between *in silico* and observed T-RF sizes might occur which makes the T-RF assignment difficult (Schütte et al. 2008). Another disadvantage of T-RFLP is the occurrence of

pseudo T-RFs which leads to an overestimation of the diversity (Egert and Friedrich 2003). In order to improve the reproducibility and accuracy of the method and the taxonomic assignment of T-RFs several technical improvements have been suggested, eg. the use of more than one restriction enzyme (Osborne et al. 2006; Padmasiri et al. 2007; Collins et al. 2003), an appropriate choice of restriction enzymes and primers (Schütte et al. 2008) and labeling both forward and reverse primers with two different fluorochromes (Collins et al. 2003). Technical improvements have also been proposed in the last years to improve data analysis (for more detailed information see Nocker et al. 2007; Schütte et al. 2008; Prakash et al. 2014).

Even though T-RFLP is a very simple method, data analysis and taxonomic T-RF assignment might require an additional effort. Several steps are needed before the data can be statistically analyzed. First, a fluorescence threshold needs to be established to remove baseline noise. Second, standardization between samples is needed to avoid the influence of variations in DNA concentration between samples. Finally, size binning is necessary between samples to be able to compare them. Further analysis is needed using database information or clone sequence *in silico* restriction for taxonomic T-RF assignment. As a consequence of the time-consuming analysis of T-RFLP data, several web-based tools have been developed to aid in T-RF assignment and identify plausible members of a microbial community based on T-RFLP data. These tools are also performing additional tasks such as profile comparison, statistical analysis of data and representation of similarity in the form of a dendrogram. Some examples are TAP-TRFLP (http://rdp.cme.msu.edu), torast (http://www.torast.de), MiCA (http://mica.ibest. uidaho.edu/) and T-RFPred (http://nodens.ceab.csic.es/trfpred/), with phylogenetic assignment tools (PAT) likeT-Align (http://inismor.ucd.ie/ *talign/), ARB software integrated tool, TRF-CUT (http://www.mpi-marburg.mpg.de/braker/trfcut.zip), TRiFLe (http://cegg.unige.ch/trifle/trifle.jnlp), and with T-RFLP statistical data analysis software (http://www. ibest.uidaho.edu/tool/T-RFLP_stats/index.php). Afterwards, statistical analysis can be performed with standard softwares in order to interpret the results.

T-RFLP has been extensively used to monitor structural changes and dynamics of *Archaea* and *Bacteria* in anaerobic reactors (Collins et al. 2003; Carballa et al. 2011; Pycke et al. 2011; Feng et al. 2010; Pervin et al. 2013; Zhang et al. 2014; Whang et al. 2015). In general, the community is monitored in order to relate changes in the microbial community by modifying one or several operational parameters. Bacterial and archaeal communities can be monitored in lab scale or full scale reactors and different statistical tools might be used to interpret T-RFLP results. In this sense, Blasco et al. (2014) studied changes in bacterial and archaeal communities by only using T-RFLP in two lab scale stirred tank reactors fed with autoclaved and untreated waste. Strong statistical analysis of the data was performed including PCoA, ANOVA and Spearman's rank correlation coefficient. The authors were able to determine that autoclaving as a pretreatment as

well as change of OLR influenced the microbial community structure, especially the bacterial one. Another investigation used T-RFLP to study the microbial community richness, dynamics, and organization of four full-scale anaerobic digesters during a time-course study of 45 days (Pycke et al. 2012). In this case the authors calculated the parameters established by Marzoratti et al. (2008) (Richness (Rr), Functional organization (Fo), Dynamics (Dy)) and performed cluster analysis. The authors demonstrated that the temperature has a strong effect on both bacterial and archaeal communities. The dynamics of change was very high and varied for both *Bacteria* and *Archaea* at a rate of change between 20–50% per 15 days. Moreover, the community organization of the *Bacteria* changed more in the thermophilic reactors, compared with the mesophilic ones and the *Archaea* community structure was more stable.

Several authors have used T-RFLP to determine the community structure and dynamics of methanogens, one of the key populations of the anaerobic process in anaerobic digesters (McHugh et al. 2004; Padmasiri et al. 2007; Ziganshin et al. 2013). By choosing the correct restriction enzymes it is possible to study the relation between hydrogenotrophic and acetoclastic methanogens. Padmasiri et al. (2007) were able to determine that hydrogen utilizing methanogens increased in abundance during a period of poor reactor performances when studying the effect of the degree of shear to which the biomass in an AnMBR is exposed. The authors used three restriction enzymes and were able to study the dynamics of *Methanomicrobiales* and *Methanobacteriales* (hydrogenotrophic) and *Methanosaetaceae* and *Methanosarcinales* (aceticlastic) without performing a cloning library for T-RF assignment. When combining T-RFLP (using one enzyme but forward and reverse primers labeled) with cloning and sequencing a more exact determination of the structure, diversity, dynamics and composition of the methanogenic community was achieved. Xu et al. (2010) studied the effect of two methanogenic inhibitors, BES and $CHCl_3$, on the methanogenic community structure present in an anaerobic sludge. The phylogenetic analysis of the archaeal sequences obtained from the clone library allowed determining the archaeal community composition which consisted of *Methanosaetaceae*, *Methanomicrobiales*, and *Methanobacteriales* and of yet uncultured archaeal lineages such as RC-III. Correspondingly, six fragments were detected as major peaks in the T-RFLP profiles. *In silico* digestion of the representative clones allowed to assign three T-RFs to *Methanomicrobiales*, one T-RF to *Methanobacteriales* and one T-RF to *Methanosaetaceae*. Finally, one T-RF was associated with two lineages, which could be differentiated by forward digestion. By using this strategy the authors were able to determine that acetoclastic methanogens were more sensitive to inhibitors than hydrogenotrophic methanogens as the relative abundance of *Methanosaetaceae* decreased compared to the control experiment and the replacement of acetoclastic methanogens was more pronounced in the $CHCl_3$ versus the BES incubations.

Another interesting option used to study methanogens is to perform T-RFLP profiles on the methyl coenzyme M reductase gene (*mcrA*) (Zhang et al. 2014; Ma et al. 2013; Ács et al. 2013). One recent example revealed which methanogens were involved in tetramethylammonium hydroxide (TMAH) degradation in three full scale bioreactors (Whang et al. 2015).

T-RFLP has also been used to monitor microbial communities present in lab scale hydrogen producing reactors, combined with sequencing and strain isolation, and it has been shown that a mixed microbial community developed while non hydrogen producing bacteria were also present in the reactors (Castelló et al. 2009; Perna et al. 2013; Ferraz et al. 2014). For example, when using raw cheese whey as substrate a low hydrogen yield was obtained which could be explained by the selection of a mixed fermentative population with the presence of hydrogen-producing organisms (*Clostridium*, *Ruminococcus* and *Enterobacter*) and other non-hydrogen producing fermenters (*Lactobacillus*, *Dialister* and *Prevotella*) in the same reactor (Castelló et al. 2011). Knowing how these organisms outcompete might be important in order to further optimize hydrogen production in UASB reactors. These examples demonstrate that T-RFLP has become a rapid, inexpensive, sensitive, robust and reproducible technique for the study of bacterial and archaeal communities in anaerobic bioreactors.

4.1.3 SSCP fingerprinting as a rapid and accurate molecular tool for monitoring microbial dynamics in anaerobic bioprocesses

The SSCP-based analysis relies on the propensity for single-strand DNA (ssDNA) to take a folded secondary structure in non-denaturing conditions. The secondary structure is determined by the nucleotide sequence and intramolecular interactions. Therefore, the SSCP method involves a step of heat denaturation of the PCR products where the ssDNA is formed. An electrophoresis is then performed under non-denaturing conditions, where ssDNA fragments are separated into different bands according to their difference in electrophoretic motility (Lee et al. 1996; Delbès et al. 2001). Nowadays, the procedure was adapted to an automated capillary DNA sequencer and is so-called Capillary Electrophoresis-SSCP (CE-SSCP) (Zumstein et al. 2000). An internal SSCP ladder is added in each sample to compare data and eliminate capillary-to-capillary or run-to-run variability, which has been a real technical improvement as compared to the original SSCP analysis (Quéméneur et al. 2011b). Since automation of the CE-SSCP electrophoresis allows a rapid and easy process of a maximum of 96 samples at a time, CE-SSCP fingerprinting is particularly useful to study the dynamics of the microbial populations over a long time series (Zumstein et al. 2000). Considering this, further statistical analysis can be carried out. Usually, CE-SSCP are processed and computed either with the

graphical interface Stat-Fingerprints program working under the R software environment (Michelland et al. 2009) or the SAFUM toolbox written in MATLAB (Zemb et al. 2007).

The V3 variable region of the 16S rRNA gene is one of the most common target for characterizing archaeal and bacterial communities in anaerobic digesters (Zumstein et al. 2000; Leclerc et al. 2001; Leclerc et al. 2004; Mnif et al. 2012; Jáuregui-Jáuregui et al. 2014). In literature, SSCP analysis is widely reported as useful to compare microbial diversity in response to changes in environmental or process operation conditions especially when the diversity is low (Leclerc et al. 2004; Ye et al. 2007). For instance, the archaeal diversity in anaerobic digesters is often limited to a dozen of peaks, and diversity can be rapidly characterised by simple peak counting on SSCP profiles (Leclerc et al. 2004). Bacterial diversity may also be monitored by SSCP in anaerobic ecosystems that are simplified by applying stringent conditions, *e.g.* in hydrogen-producing bioreactors where only few major bacterial species dominate and are mostly affiliated to the *Clostridium* genus (Rafrafi et al. 2013). Similarly, the CE-SSCP analysis is perfectly adapted to monitor the decrease of microbial diversity all along an enrichment procedure, *e.g.* from anaerobic sludge to a co-culture of anaerobic syntrophs (Trably et al. 2008). The diversity of more complex bacterial community can also be estimated but only when considering the background signal corresponding to the information that is not enclosed in peaks areas (Loisel et al. 2006).

Interestingly, Bauer et al. (2008) developed and validated successfully a functional PCR-SSCP method followed by a cloning step of the interesting bands and a sequence analysis. They evaluated this method in combination with direct PCR cloning and sequence analysis. A novel and highly degenerated *mcrA/mrtA* pair of PCR-primers targeting all known methanogens was developed and tested with reference methanogens. DNA extracts from biogas fermenters fed with maize silage and operated at different conditions were analyzed and the dominant residing methanogens were characterized to assess their functional importance during the process stages. Whereas direct PCR cloning was more suited to determine the relative abundances of methanogens, the SSCP analysis easily detected population changes and the involved sequences specifically. Although most of the bands were not sharp probably because of high degenerated primers, the authors concluded that the SSCP technique was more suitable than DGGE for monitoring methanogens. The SSCP method was further refined and employed to investigate the presence and dynamics of distinct sub-populations of methanogens during biogas process acidification (Munk et al. 2010). The authors concluded that technical details associated with degeneration of the primers should be improved since different sequences comigrated in one SSCP band and identical sequences were found in different bands. This method, in combination with classical chemical parameters, could help to quickly indicate changes in population structure due to process conditions, and thus prevent dysfunction of the process.

The correlation between the expression of *hydA* gene and H_2 production in batch or continuous reactors has been well established (Tolvanen et al. 2008a; Wang et al. 2008a; Tolvanen et al. 2008b; Wang et al. 2008b). Quéméneur et al. (2010) developed a successful CE-SSCP method based on functional *hydA* [Fe-Fe]-hydrogenase genes for monitoring hydrogen (H_2)-producing clostridial populations. These authors designed and validated a set of non-degenerated primers, named *hydAClosF/hydAClosR*, by binding two conserved regions of the active site domain (H-cluster) of the [Fe-Fe]-H_2ases-coding genes. These primers showed a high *in silico* specificity and selectivity for the detection of the *hydA* genes from a broad range of *Clostridium* strains belonging to different phylogenetic clusters. This method was validated using experimental H_2-producing mixed cultures and was shown to be very effective in monitoring the shift of H_2-producing bacterial populations in relation to environmental changes, such as temperature (Quéméneur et al. 2010), pH (Quéméneur et al. 2011b) and the type of carbohydrate as substrate (Quéméneur et al. 2011a). The *hydA* sequences amplified with this new set of primers clearly showed that the functional diversity of *hydA*-carrying *Clostridium* sp. strains impacted the H_2 production yields. Interestingly, the 16S rDNA-based fingerprints were found to be less correlated to changes in H_2 production performances. The difficulty in distinguishing *Clostridium* cluster I species in H_2 dark fermentation systems by using universal bacterial primers was also reported by (Chang et al. 2008).

Hence, functional CE-SSCP tools provide highly reliable and throughput analysis of the microbial communities which could be used as a complement to 16S rDNA phylogenetic markers. Such first successful approach for determining functional microbial dynamics in bioprocesses by the CE-SSCP technique makes possible its rapid extent to other functional groups found in anaerobic digestion systems, e.g. methanotrophic (*pmoA*), and sulfate reducing (*dsrAB*) bacteria.

The three fingerprinting methods presented here are highly suitable to study the microbial diversity, structure and dynamics of microbial communities in AD reactors. Overall, the CE-SSCP method offers several advantages in terms of rapidity and ease of obtaining data, as well as a high reproducibility between runs and data processing. Regarding the DGGE-based methods, which relies on the use of clamped primers and gradient gels, with only a limited number of samples per gel that are analysed, SSCP and T-RFLP are simpler and more straight forward techniques (Marzorati et al. 2008; Rastogi and Sani 2011, Schütte et al. 2008). However, DGGE is widely used and has turned into a routine fingerprinting method to study microbial diversity because it has reasonable resolving power and allows fast comparative analysis between different samples (Van Elsas and Boersma 2011). On the other hand, T-RFLP has been shown to be a reproducible and accurate tool for community fingerprinting (Osborn et

al. 2000) but, the data analysis requires an additional effort. Additionally, even though DGGE is widely used, special equipment, but very affordable, is required while for SSCP and T-RFLP the samples can be analysed in any sequencing service commercially available.

5 – Tools for microbial quantification (How many microorganisms from the different groups are present in the reactor?)

5.1 Quantification and detection of particular organisms in anaerobic digesters by FISH

The FISH method involves application of fluorescently labeled probes to ribosomal rRNA in permeabilized whole microbial cells. The probes consist of short pieces of DNA (usually 15-25 nucleotides in length) that are designed to specifically hybridize to their complementary target sequence on the rRNA structures (16S and 23S subunits are typically used for *Bacteria* and *Archaea*) (Amann et al. 1995). From the composition of the probe, it is possible to design it to specifically target a narrow phylogenetic group (down to the species level) or any other higher phylogenetic hierarchical group (Amman et al. 1995). No probes will hybridize to those cells without target sequences. Cells containing the target sequence will on the other hand retain the hybridized probe and due to the large number of ribosomes only the active cells (from 10^3 to 10^5ribosomes) become fluorescently labeled. FISH provides an effective means of identification and qualitative and/or quantitative microbial population analysis in natural and engineered environments (Amann et al. 2001).

The optimisation process in FISH method is generally carried out with pure cultures of microorganisms that contain rRNA with a perfect match to the probes (positive controls) and with microorganisms that contain rRNA with preferably one (optimally centrally located) or several mismatches to the probe (negative controls) (Manz et al. 1992). There are several variables that can be exploited to modify the strength of a hybrid between an oligonucleotide and its perfect target. These include temperature, ionic strength of the medium and organic solvent concentration (Amann et al. 1995).

Several researches applied FISH in order to study microbial ecology in aerobic and/or anaerobic reactors (Amann et al. 1992; Raskin et al. 1994; Power et al. 1999; Rocheleau et al. 1999; Sekiguchi et al. 1999; Araujo et al. 2000; Araujo et al. 2004; Imachi et al. 2000; Zheng and Raskin 2000; Egli et al. 2003; Ginige et al. 2004; Díaz et al. 2006). One of the early studies (Raskin et al. 1994) is very important to mention because authors designed different oligonucleotide probes for ex situ hybridization in order to quantify different groups of methanogenic *Archaea*. The second study, which is also worth to mention, was of Sekiguchi et al. (1999). This was the first study that used FISH combined with confocal laser scanning

755 microscopy (CLSM) to elucidate the spatial distribution of microbes within two types of
756 methanogenic granular sludge (mesophilic and thermophilic) in UASB reactors fed with
757 sucrose-, acetate-, and propionate-based artificial wastewater. In situ hybridization with
758 archaeal- and bacterial-domain probes within granule sections clearly showed that both
759 mesophilic and thermophilic granules had layered structures and that the outer layer harbored
760 mainlybacterial cells while the inner layer consisted mainly of archaeal cells. *Methanosaeta*-,
761 *Methanobacterium*-, *Methanospirillum*-, and *Methanosarcina*-like cells were detected with
762 oligonucleotide probes specific for the different groups of methanogens, and they were found to
763 be localized inside the granules, in both types of which dominant methanogens were members
764 of the genus *Methanosaeta.*

765 In the study of Crocetti et al. (2006), 3000 Euryarchaeota 16S rRNA gene sequences were
766 phylogenetically analyzed and previously published oligonucleotide probes were evaluated for
767 target accuracy. Where necessary, modifications were introduced or new probes were designed.

768 Despite all the studies mentioned above that used FISH, there are some limitations associated
769 with this method, such as: 1-poor cell permeability (a pretreatment is necessary to fix Gram
770 positive cells or some archaea (Davenport et al. 2000; Zheng and Raskin 2000). 2-The detection
771 limit is not very low (varies from 10^3 cells/mL (Li and Gu 2011) to 10^4 cells /mL (Amann
772 1995). 3-Insufficient ribosome content, as inactive or dormant cells will not give fluorescent
773 signal or it can be very low (Morgenroth et al. 2000). 4- Sample autofluorescence, specially
774 some methanogenic *Archaea* can produce autofluorescence that interfere with fluorescent
775 labels.

776 Some of these limitations, especially the detection of cells with low ribosomal content present in
777 oligotrophic conditions, can be overcome by applying the catalyzed reporter deposition FISH
778 (CARD-FISH) (Pernthaler at al. 2002). In this approach (CARD-FISH) the critical step is the
779 diffusion of the large molecules, in this case the horse radish peroxidase (HRP)-labelled probe,
780 into whole cells embedded in an agarose matrix. A directed permeabilization of prokaryotic cell
781 walls prior to the hybridization step is of crucial importance to enable the penetration of the
782 probe (Pernthaler et al. 2002). In the study of Wilhartitz et al. (2007), CARD-FISH resulted in
783 substantially higher recovery efficiency than the conventional FISH-approach and therefore it is
784 more suitable for the enumeration of specific prokaryotic groups in ultra-oligotrophic
785 environments such as ground- and drinking water samples.

786 Therefore, knowledge of the nature and applicability of the sample as a uniform protocol with
787 application of the proper controls are of fundamental importance for obtaining solid and
788 comparable information. By using FISH it is possible to observe the morphology and to
789 quantify numbers of bacteria or equivalent biovolume. Although the method is very time

consuming and tedious, the principal advantage is that it is possible to quantify and at the same time, determine the position of the different cells in the community.

5.2 Q-PCR for quantification in anaerobic digestion processes

The quantitative polymerase chain reaction (q-PCR) or real-time quantitative PCR (RT-PCR) is a molecular technique that detects and quantifies DNA molecules in solution, by means of fluorescence signal detection. Based on the end point or regular PCR, a target region of the template DNA is amplified with specific primers, although a key difference is that products are detected in each cycle during the exponential phase of the process (Higuchi et al. 1992).

The two detection methods most widely used in environmental microbiology are: i) the addition of intercalating fluorescent dyes (e.g. SYBR Green I), which emits fluorescence only when binds to double-stranded DNA (Wittwer et al. 1997) or, ii) sequence-specific oligonucleotide probes (complementary to an internal sequence) labeled with a fluorophore and with a quencher (e.g.TaqMan) (Gudnason et al. 2007). The TaqMan technology (Heid et al. 1996) is quite more specific than SYBR Green I due to the third oligonucleotide probe used in the reaction that drastically reduces false positives. On the other hand, SYBR Green I is a less expensive method and sometimes the only choice, when the target sequence hinders the primer and probe design.

Hence, the quantification of the fluorescent signal can be correlated to the amount of target DNA molecules present in the solution. Quantification can be absolute or relative; the former method is currently the most used in microbial ecology and consists in the interpolation of the inquired samples in a standard curve constructed with known concentrations of target sequences, often cloned in plasmids or as PCR products.

The specificity, sensitivity but mainly the quantitative basis of the q-PCR filled a gap in microbial ecology. Since the pioneer studies in 2000s (Suzuki et al. 2000; Takai and Horikoshi 2000), this approach has been increasingly applied to investigate microbial ecology questions that could not be addressed with other techniques such as clone libraries or even the actual high-throughput DNA sequencing. It has been used to track genes and/or transcripts of ribosomal or functional genes across temporal or spatial scales in different environments such as biofilms (de Gregoris et al. 2011), rumen (Bekele et al. 2011) seawater (Mincer et al. 2007) and alpine soils (Brankatschk et al. 2011). A recent work published by Kim et al. (2013) describes the state of the art in the quantification of the different groups of microorganisms found in wastewater treatment systems.

In the anaerobic wastewater treatment research field, q-PCR has been applied mainly to monitor the methanogen community. The strategy consists in targeting the 16S rRNA gene (Yu et al. 2005) or the *mcrA* gene (Alvarado et al. 2014). By targeting the 16S rRNA gene, some studies

have shown a swhich from aceticlastic to hydrogenotrophic methanogenesis when stressing conditions such as temperature or ammonia are applied to lab-scale reactors (Lee et al. 2009; Mc Keown et al. 2009; Song et al. 2010; Traversi et al. 2011; Jang et al. 2014; Town et al. 2014). On the other hand, a direct correlation between specific methanogenic activity and *mcrA* gene copy numbers was reported recently (Morris et al. 2013).

Some considerations must be stressed before interpreting or comparing q-PCR results. The first issue to consider is the DNA yield bias introduced by the different methods of nucleic acids extraction (Martin-Laurent et al. 2001). Another important aspect is the operon copy number variation (CNV) of the target gene, in the case of bacterial 16S rRNA, variations of 1-15 copies between different species or even in different phases of the cell cycle were reported (Klappenbach et al. 2000; Sun et al. 2013). In the case of organisms from the *Archaea* domain, ribosomal operon variation is less pronounced, ranging 1-4 copies per genome (Kim et al. 2013). As an attempt to cope with the gene CNV, alternative molecular markers with only one copy per genome have been used. The *mcrA* gene, mentioned above has been proved to be a good molecular marker for tracking methanogens in the archaeal community (Luton et al. 2002).

The specificity of the primers and probes used is another key aspect of this approach. However, there is a vast set of primers designed to monitor the wastewater treatment microbiology. For instance, a complete review of primers and probes designed to monitor methanogenic *Archaea* have been compiled by Narihiro and Sekiguchi (2011).

Due to the potential applications of this technique in wastewater treatment research, efforts must be done to jointly improve the robustness of the approach and standardize protocols to compare results between different works. Although far from a golden standard, q-PCR is currently a promising technique to assess quantitatively the structure of microbial communities. The principal advantages of this quantification technique are the low detection limits (it is possible to detect a single DNA molecule depending of the efficiency of the reaction), the possibility to perform several samples at the same time and the short time required to obtain results. q-PCR analysis give access to absolute quantification, and thus complement very well sequencing and FISH data that give relative abundances of targeted microorganisms.

6. Tools to analyze microbial function (What is the function of the microorganisms inside the reactor?)

Understanding how complex microbial communities function inside anaerobic digesters and how environmental conditions may affect interspecies relationships is currently one of the big challenges for both environmental biotechnologists and microbiologists. Only by getting more

insight into microbial function we will be able to fully exploit and manage microbial communities for biotechnology purposes (Verstraete 2007). Over the last years, several culture-dependent and culture-independent techniques have been developed that allow to study community composition and function in complex ecosystems, such as detection of specific enzyme activity or of specific enzyme-coding genes, functional microarrays, radioactive and stable isotope labeling, probing and advanced imaging, and meta-omics approaches (Su et al. 2012; Vanwonterghem et al. 2014). These techniques, alone or in combination with each other and with differential bioreactor operation, can offer insights into the metabolic activities of microbial communities in AD processes and it is in this perspective that they are discussed in the following sections.

6.1 Stable isotope probing and in situ advanced microscopic methods

Stable isotope probing (SIP) is performed by amending a stable isotope (for example ^{13}C-labeled substrate to microbial communities and further analyzing microbial composition in heavy DNA or rRNA fraction, *e.g.* by cloning and sequencing (Figure 4a) (Radajewski et al. 2000). Heavy DNA/rRNA represents the active populations in the community because those are the ones incorporating the isotopes within their biomass.

In recent years, with advances in high resolution imaging and spectroscopy, SIP has been complemented with other techniques, such as FISH, microautoradiography combined with FISH (MAR-FISH), Raman microspectroscopy combined with FISH (Raman-FISH), or nanoscale secondary ion mass spectrometry (NanoSIMS) combined with FISH (SIMSISH) (Chapleur et al. 2013; Huang et al. 2007; Lee et al. 1999; Li et al. 2008). These techniques allow to study taxonomic identity and activity of microbial communities at single cell resolution (Figure 4b). This is a field of fast evolution and complementary techniques are rapidly emerging. One example (not yet applied to anaerobic digestion systems) is the utilization of amino acid tagging and click chemistry for *in situ* visualization of newly synthesized proteins (biorthogonal non-canonical amino acid tagging, BONCAT) (Hatzenpichler et al. 2014). Combination of BONCAT technique with FISH (BONCAT-FISH) enables a direct link between taxonomic identity and translational activity. Comparison of BONCAT-FISH results and SIMISH for anabolic nitrogen assimilation showed concordance, with BONCAT-FISH offering a less expensive solution for these studies.

Several studies have addressed the key players in anaerobic digesters by performing separate feeding studies with different ^{13}C-labelled carbon sources, such as ^{13}C-cellulose, ^{13}C-glucose, ^{13}C-acetate and ^{13}C-propionate, followed by DNA-SIP and FISH analysis (a summary of the main results obtained in these studies is shown in Table 2). In this way, the metabolic activity

for independent steps in the cellulose/glucose to methane conversion pathway in methanogenic reactors could be assigned to specific taxonomic groups, *e.g.* cellulose degraders were shown to belong to *Clostridia* and *Acetivibrio*, while subsequent glucose conversion was mainly performed by bacteria belonging to the *Clostridia* and *Porphyromonadacaea* (Li et al. 2009; Li et al. 2008). The involvement in cellulose conversion of novel identified groups in anaerobic digesters, such as the WWE1 candidate division, could also be confirmed using these methods (Limam et al. 2014). The strategy to perform feeding studies with different ^{13}C-labelled substrates, corresponding to the different levels of the conversion pathway, was also successfully applied in elucidating the microbial groups involved in each of the sequential steps. In a glucose-degrading anaerobic digester it was elegantly demonstrated that glucose was converted to propionate by members belonging to *Actinobacteria*, *Bacteroidetes* and *Chloroflexi*, and that propionate was further converted to acetate by *Syntrophobacter* and *Smithella*, and acetate finally utilized by members belonging to the *Synergistes* group 4 and the methanogen *Methanosaeta* (Ito et al. 2011; Ito et al. 2012).

6.2 The meta-omics era

Omics techniques (genomics, transcriptomics, proteomics) were first used to study and characterize cultivable isolates but, with the development of more powerful instrumentation and bioinformatics tools, application of these techniques has been extended to more complex ecosystems. This is important because interspecies interaction and responses to environmental conditions, which result in multiple community behaviors, can only be identified if the ecosystem is assessed as a whole. Meta-omics techniques - metagenomics, metatranscriptomics, metaproteomics, and metabolomics - can be applied to ecosystems to unravel composition and function of uncultured microorganisms (Figure 5). Methods such as SIP and advanced imaging techniques discussed in the previous section allow analyzing structure-function in ecosystems, but only associated with a particular physiological trait. Omics techniques, on the other hand, have the potential to reveal the complete picture of the microbial functionalities in an ecosystem.

6. 2. 1 Metagenomics

Metagenomics has greatly facilitated the mining of microbial diversity and metabolic potential of highly diverse microbial communities. In particular, in recent years the advances in NGS technology and accompanying bioinformatic tools have enabled the generation of large sequence datasets and streamlined analysis of the complex sequence data. Through these methods, in-depth 16S rRNA gene-sequencing analysis of production-scale biogas digesters has

generated many novel details on the microbial composition (and compositional changes) during anaerobic digestion. These studies confirmed that *Clostridia* comprise the most prevalent bacterial class in different types of AD reactors (Jaenicke et al. 2011; Li et al. 2013; Rademacher et al. 2012; Schlüter et al. 2008), but also identified hitherto neglected taxa such as *Streptococcus, Acetivibrio, Garciella, Tissierella, Gelria* (Jaenicke et al. 2011) or *Psychrobacter* and *Anaerococcus* (Li et al. 2013). Methanogenic archaea were mainly represented by the hydrogenotrophic *Methanoculleus, Methanosarcina, Methanobacterium* and *Methanothermobacter*, or the acetoclastic archaea *Methanosaeta*. This corresponded to results from previous studies. In addition one study also demonstrated the presence of *Thermacetogenium* (Rademacher et al. 2012). The type strain of this genus, *Thermacetogenium phaeum*, oxidizes acetate with *Methanothermobacter thermoautotrophicus* (Hattori et al. 2000).

Besides phylogenetic information, the metagenomics analysis also provides more details on the functional potential of the microbial community and its correlation to the anaerobic process under study. In AD reactors treating plant material waste a high number of sequence reads related to cellulose and hemicellulose conversion as well as other carbohydrate-degrading genes were detected, which in one study could be assigned to the predominant phylogenetic taxa identified (Jaenicke et al. 2011). *Acetivibrio* species had previously been recovered in the heavy DNA-fraction during SIP experiments with cellulose (Li et al. 2009); taxonomic and functional genes results obtained by Jaenicke et al. (2011) support once more that this genus most probably play a role in cellulose degradation. Some Clostridia may be involved in acetogenesis, as deduced by mapping bacteria taxa to metagenome hits representing the Wood-Ljungdahl pathway (Jaenicke et al. 2011). Reads encoding for enzymes required for hydrogenotrophic and acetoclastic methanogenesis pathways were also detected, corresponding to the taxonomic observations (Figure 6). Enzymes necessary for catalyzing CO_2 conversion to methane were all found within the metagenomes of the bioreactor sludges treating cellulosic materials and waste activated sludge (Rademacher et al. 2012; Wong et al. 2013). Acetate kinase and phosphotransacetylase, involved in the first step of acetoclastic methanogenesis, were not detected in these studies. Wong et al. (2013) proposed an alternative route in which an acetate transporter coupled to acetyl-CoA synthetase and the hydrolysis of pyrophosphate by inorganic pyrophosphatase drives this reaction forward, similar to the pathway present in *Methanosaeta thermophila* (Smith and Ingram-Smith 2007).

One of the main drawbacks of metagenomics studies is still the large amount of sequence reads with unidentified microbial origin or gene prediction. This may indicate that the anaerobic digestion process is also conducted by yet unknown microorganisms or species for which no genomic information is available. Most metagenomics studies rely on complete or draft genomes to identify fragmentary sequences, which limits the ability to resolve metagenomics

data deriving from unknown microbial diversity (Temperton and Giovannoni 2012). Therefore, it is important to continue the effort of isolating new microorganisms and sequencing the genome of representatives of the different groups of microorganisms present in bioreactors, following a similar strategy that was successfully applied with the human microbiote.

6. 2. 2 Metatranscriptomics

While metagenomics gives information on the metabolic and functional potential of a microbial community, it cannot provide information on the activity of the genes present or to differentiate between expressed and non-expressed genes. To get an estimate of the actual metabolic activity, total mRNA from a microbial environment needs to be retrieved and sequenced. Metatranscriptomics in general generates less complex datasets to analyze (only transcribed genes are retrieved) than metagenomics does. However, mRNA molecules have a short half-life, and ribosomal RNA represents the majority of RNA isolated, thereby reducing sequencing-depth for mRNA reads that represent the expressed genes and pathways.

Recently, a metatranscriptome study investigated the transcriptionally active microbial fraction of a biogas-producing reactor and identified the most abundant transcripts during biogas production (Zakrzewski et al. 2012). This is the only metatranscriptomics study performed on a full-scale anaerobic digester thus far. The most abundant mRNA-derived sequence reads from this metatranscriptome dataset were from transcripts encoding enzymes involved in substrate hydrolysis, acidogenesis, acetate formation and methanogenesis, indicating that these pathways were highly active in the anaerobic digestion process. Taxonomic profiling of the active community, based on 16S rRNA tags, revealed that members of the *Firmicutes* and *Euryarchaeota* were the most abundant bacteria and archaeal phyla in the system, respectively; low numbers of 16S rRNA tags were associated to the phyla *Bacteroidetes*, *Actinobacteria*, and *Synergistetes*. In addition, comparison of the metatranscriptome dataset with 16S rRNA-gene metagenomic sequence tags derived from the same community, indicated that the archaeal species present were highly active and that in general the most abundant species in the community contributed to the majority of the retrieved transcripts.

6. 2. 3 Metaproteomics

Where metagenomics identifies the microbial community and its metabolic/functional capacity, and metatranscriptomics the transcriptional activity of specific metabolic pathways/microbes, metaproteomics can be used to characterize highly expressed proteins within an environmental microbial consortium. This can provide a more functional evidence of key metabolic pathways that are active and important for anaerobic digestion processes. In contrast to metagenomics and

metatranscriptomics, the main limitation of metaproteomics is to extract a sufficient quantity of high quality proteins that are representative of the sample. Extracting a representative protein fraction for analysis is complicated due to the complex microbial communities involved, the presence of interfering compounds, and the heterogeneity of natural environments. Nevertheless, metaproteomics has a great potential to link the genetic diversity and activities of microbial communities with their impact on ecosystem function (Su et al. 2012).

Metaproteomics studies in anaerobic digestion are still limited. One study analysing a lab-scale anaerobic bioreactor treating glucose-based wastewater assigned putative functions to the proteins detected, and demonstrated that they were mostly related to cellular processes such as methanogenesis from both CO_2 and acetate, glycolysis and the pentose phosphate pathway (Abram et al. 2011). The number of identified proteins was very low though, only 18 distinct proteins were assigned to different functions (from 70 proteins excised from two-dimensional electrophoresis gel). No metagenomics data was available from the reactor, which decreases the rate of identification. The protein assignment also indicated the presence of specific microorganisms in the bioreactor (*Actinobacteria*, *Firmicutes*, *Proteobacteria*, *Methanosarcinales* and *Methanobacteriales*), though little overlap was observed with the 16S rRNA gene clone libraries constructed for phylogenetic analysis of the bioreactor. In a similar work, Hanreich et al. (2013) analyzed the metaproteome of biogas batch fermentation systems co-digesting straw and hay. Members of the orders *Clostridiales*, *Bacteroidales* and *Flavobacteriales* were identified as key players in this ecosystem from the analysis of metagenomics data. A higher number of gene sequences codifying for enzymes from the glycoside hydrolase families GH5, GH48 and GH94 (comprising cellulase, cellobiohydrolyse and cellobiose phosphorylase activities) were present at the beginning of the incubations (together with higher numbers of *Clostridiales*-related microorganisms); over time, an increase in *Bacteroidales* and *Flavobacteriales* was observed with an increase in the number of sequences encoding enzymes of the GH2, GH13, GH29, GH77 and GH92 families (comprising enzymes capable of cleaving a variety of monosaccharides and polysaccharides). Nevertheless, only three hydrolytic enzymes could be detected during the proteomics analysis – a putative α-amylase, a pectaselyase and a glycosyide hydrolase family protein). On the other hand, a high number of enzymes involved in methanogenesis could be detected in the protein fractions (e.g. methyl-coenzyme M reductase), although methanogens represented a minor part of the community as inferred from metagenomic data (orders *Methanosarcinales*, *Methanomicrobiales*, and *Methanobacteriales*). Recently, Lü et al. (2014) performed a much more comprehensive metaproteomics study on thermophilic anaerobic cellulose degradation, with the identification of more than 500 non-redundant protein functions. The taxonomic distribution of the identified proteins suggested the presence of a few dominant bacterial groups,

namely *Caldicellulosiruptor* species (≈ 41%), *Coprothermobacter proteolyticus* (≈ 20%) and *Clostridium thermocellum* (≈ 16%). *Methanothermobacter thermoautotrophicus* (≈ 5%) was the most abundant methanogen. A large number of proteins potentially involved in cellulose and hemicellulose binding and hydrolysis or in oligosaccharide metabolism were identified, and their synthesis linked to bacteria related to *Caldicellulosiruptor* spp. and *Clostridium thermocellum*. A cooperation between these two groups of bacteria was suggested, in which *Caldicellulosiruptor* spp. are responsible for hydrolysing the cellulosic part of the substrate, whereas *Clostridium thermocellum* is also hydrolysing hemicellulose in addition to cellulose. Regarding methanogenesis, data generated by pyrosequencing, metaproteomics and isotopic analyses together strongly support the predominance of the hydrogenotrophic pathway by strains of the genus *Methanothermobacter*. None of the enzymes specific for acetoclastic methanogenesis were identified, which opens the hypothesis of the occurrence of syntrophic acetate oxidation (SAO) in this ecosystem. Nevertheless, the lack of known protein specific for the SAO pathways and the limited knowledge about thermophilic SAO microorganisms did not allow further confirming this. SAO is currently a topic of interest as new microorganisms and theories for this pathway are emerging (Nobu et al. 2015).

7- Statistical analysis to link results from molecular data to reactors performance

The results obtained by molecular methods can be linked to reactor performances using different statistical analysis. The link with environmental data is very important for the interpretation of the results. Although some conclusions can be drawn just by observing fingerprinting profiles or OTU tables, a deeper analysis is essential in order to arrive at strong and robust conclusions.

There are different kinds of analyses that can be performed with different levels of complexity. The most frequently used are the hierarchical clustering analysis, the ordination methods and the calculation of diversity indices (Talbot et al. 2008). The methods can be applied to either fingerprinting data or sequencing analysis datasets. The first step is to perform a matrix with the samples and the presence/absence or abundance of each "species" (represented by OTUs, chromatogram peaks or DGGE bands). With this species matrix several analyses can be performed.

7.1 Multivariable statistical tools

With the hierarchical cluster analysis the samples are grouped according to the dissimilarity between them, the results are visualized in a dendrogram constructed to represent how the samples are grouped and different coefficients and algorithms can be applied for the clustering

(Ramette 2007). This analysis is very useful to determine the changes in the communities due to differences in operation conditions. The samples can also be ordinated according to the dissimilarities between them in two dimensional spaces using ordination methods as Principal Component Analysis (PCA), Non-metric Multidimensional Scaling (NMDS) or Canonical Correspondence Analysis (CCA). In constrained (canonical) ordination analysis methods it is possible to include in the original matrix additional environmental data such as: pH, temperature, organic loading rate (OLR), or other environmental data. The resulting ordination diagrams display samples, species, and environmental variables so that 'fitted species samples' and 'species environment' relationships can be derived as easily as possible from angles between arrows or distances between points and arrows (Ramette 2007).

Another objective may be to test whether differences between groups of objects (rows) in a multivariate table are significantly different based on the set of their attributes (columns), i.e. to test whether similarities within groups are higher than those between groups. Here, nonparametric multivariate ANOVA (NPMANOVA) and analysis of similarities (ANOSIM), which are commonly found in standard statistical packages, can be used (Ramette 2007).

An overview of the analysis that can be applied in microbial ecology could be found in the review published by Ramette (2007) and in the web tool developed by Buttigieg and Ramette (2014).

7.2 Diversity indices

Using the same matrix of species abundance, diversity indices can be calculated. The indices reflect the organization of the community and can be used to determine differences within microbial communities (Talbot et al. 2008). Three main diversity indices: richness, evenness (or equitability), and Shannon index, can be calculated from fingerprinting or sequencing data to define the structure of a microbial community (Hill et al. 2003). Richness is simply defined as the number of different OTUs obtained by the fingerprinting method or sequencing approach. Evenness is a measure of the equitability of abundance. In other words, an environmental sample with more even abundance is more diverse than a sample with predominant and sparse species. Finally, Shannon index (H_o) is calculated from the relative abundances of each band/peak/OTU and is also influenced by the number of different OTUs. There are other diversity indices described in the bibliography that could be also determined, more detailed information can be found at the work of Hill et al. (2003).

The clustering, ordination methods and diversity indices can be calculated using different free softwares that are easily downloaded from the web (R package; http://www.r-project.org/, PAST package: http://palaeo-electronica.org/2001_1/past/issue1_01.htm, Hammer 2001).

7.3 Microbial Resource Management concept

Marzoratti et al. (2008) proposed a new concept to obtain valuable information from fingerprinting data that can be applied for the management of microbial communities for biotechnological applications. The idea of the authors was to define parameters that describe the organization of a community and reflects its functionality. These parameters can be used to explain and predict the performance of reactors.

Three parameters were proposed to describe the structure and assembly of the microbial communities: Richness (Rr), Functional organization (Fo), Dynamics (Dy).

The Rr represents the number of different individuals (species) in a sample and is calculated as the number of peaks in a fingerprinting profile (T-RFLP or SSCP), the number of bands in DGGE or the number of different OTUs in a cloning library. In the case of DGGE the authors propose a rarified coefficient to give different weight according the identity of the microorganisms represented by a band. This coefficient was calculated according to the migration of the bands in the gel.

The Fo coefficient is a single value that describes a specific degree of evenness, measuring the normalized area between a given Lorenz curve and the perfect evenness line. This coefficient can be determined by constructing Lorenz curves and calculating Gini coefficients (Rousseau et al. 1999) or by constructing Pareto-Lorenz curves (Wittebolle et al. 2008). The higher the Fo coefficient, the more uneven a community is.

The Dy is a coefficient to determine the rate of change of a community along time and is calculated from the similarity values within two samples taken in a frame of time (for example if the samples are taken weekly the frame of time is one week). The similarity between two samples is determined from the cluster analysis.

The authors propose that, according to the ecosystems functionality, the community will adapt to a particular assembly, represented by the three parameters that indicate if the community is in a “good or bad” organization to perform a function, deviation from these “normal” values may represent problems in the function (Marzorati et al. 2008).

These parameters were used to describe the communities from full scale anaerobic bioreactors (Pycke et al. 2009), to understand methanogenic bioreactors failures after overload (Carballa et al. 2009) and to correlate microbial communities with different operational conditions (Werner et al. 2011; Ciotola et al. 2013; Chen et al. 2014).

8- Lessons from the application of molecular tools to study the anaerobic digestion process and future challenges

During the last three decades molecular biology tools have been more and more applied to the study of the anaerobic digestion process. Each method present different advantages and limitation which have to be taken into account (Table 3). Important questions have been addressed and important information was generated. The key point now is how to use this information to improve the process.

From the 16S rRNA gene approach we now know that the microbial communities involved in the anaerobic digestion present a general microbial composition at the phylum level that is shared in most cases. These following phyla: *Chloroflexi*, *Firmicutes*, *Proteobacteria*, *Bacteroidetes*, *Synergistetes*, *Verrucomicrobia* predominate within the *Bacteria* Domain (Nelson et al. 2011; Werner et al. 2011; Sundberg et al. 2011). Species of known anaerobic microorganisms with important functions for the anaerobic digestion (hydrolysis, fermentation, acetogenesis) are included in these phyla. However, three of these phyla (*Chloroflexi*, *Verrucomicrobia* and *Synergistetes*) are represented by very few isolates and large amount of environmental sequences, then, there is a big hole in the knowledge of these organisms and their role in anaerobic digestion.

In particular, the organisms from the *Chloroflexi* phylum presented high abundance in the biomass from methanogenic reactors, and form part of the "core group" of organisms always found in anaerobic digesters (Riviere et al. 2009; Werner et al. 2011). It was postulated that they play an important role in the hydrolytic and fermentative step and in the formation of granules or flocs (Sekiguchi et al. 2001; Yamada et al. 2005). They were also related to bulking problems in some UASB reactors (Yamada et al. 2007; Borzacconi et al. 2008).These organisms are very difficult to culture and isolate; therefore, new isolation procedures and new molecular techniques should be applied to know the role of these organisms so abundant in AD systems.

On the other hand, it is known that the syntrophic organisms are a very specialized group of bacteria with low abundance in reactors and that they present high resilience (Werner et al. 2011). Important efforts have been made to understand the ecology of these key microorganisms (Stams et al. 2012) and there is still much work to be done in the physiology of these organisms with such a particular way of obtaining energy.

It is also known that *Archaea* from the *Euryarcheota* phylum are responsible for the methanogenic step. Within methanogens two metabolic pathways have been found: the hydrogenotrophic and acetoclastic. Although it was postulated for several years that the acetoclastic pathway predominate in the AD bioreactors (García et al. 2000; Batstone et al. 2006), the molecular data now are questioning this paradigm as the hydrogenotrophic

1175 methanogens have been found to be more abundant in several investigations (Kampmann et al.
1176 2012). This is not trivial because the different groups of methanogens present different kinetics
1177 and different tolerances to toxics compounds. Hence, there is still a need of knowledge within
1178 this important group.

1179 Although the general composition of the biomass from AD process is, in general, maintained at
1180 the phylum level, the particular species that predominate in each system change. These changes
1181 during operation and between different systems demonstrate the high redundancy within the
1182 different guilds. It was postulated that this high redundancy is one of the keys to maintain the
1183 general function of the process in a changeable environment (Fernández et al. 1999). It was also
1184 demonstrated by several authors that the AD community is very dynamic (Fernández et al.
1185 1999, Zumstein et al. 2000; Pycke et al. 2011). This high dynamic was proven at the species
1186 level (represented by OTUs, fingerprinting peaks or bands), but it is not known whether this
1187 dynamic indicates change in the metabolic pathways or not. Future works using metagenomic
1188 and metaproteomics approaches are necessary to answer this important question.

1189 Several research groups have been focused in trying to explain the environmental factors that
1190 drive the AD process; from these works, it is possible to conclude that temperature of operation
1191 (Levén et al. 2007), type of substrate (Zhang et al. 2014) and inoculum are important factors
1192 that select different populations (Ali Shah et al. 2014). This is an important issue which should
1193 be taken into account for practical applications as they could aid in decrease start-up periods of
1194 full scale reactors and predict failures due to changes during operation. There is a need of
1195 compilation of all the information to have an overview of the different variables that affect the
1196 microbial communities in AD processes and how the communities cope with the changes during
1197 the operation. In that sense, a web site where all the information on the physiology of all the
1198 organisms present in the AD process, as was performed for the activated sludge process
1199 (http://midasfieldguide.org), would be a good initiative to follow.

1200 Finally, all the information generated should be applied to try to explain the function of reactors
1201 and help to solve engineering problems. Figure 7 gives an overview of the possible application
1202 of molecular biology tools to improve AD processes. The development of new molecular
1203 techniques focusing on the function, the application of deeper statistical analysis and the
1204 improvement of bioinformatic tools to manage the high amount of data generated will help to
1205 walk in that direction. An effort to integrate the four disciplines: microbial ecology,
1206 environmental engineering, mathematical modeling and bioinformatics will be necessary to
1207 achieve this goal.

1208

1209 **Acknowledgements**

The authors would like to thanks Dr. Rosario Durán and Lic. Analía Lima Raimondo (Unidad de Bioquímica y Proteómica Analíticas, Institut Pasteur de Montevideo, Uruguay) for proteins 2D gel figure, and Eng. Magela Odriozzola (BioProa Laboratory, Instituto de Ingeniería Química, Facultad de Ingeniería, UDELAR) for providing reactor scheme. The authors thanks to the following funding agencies and projects, ANII FSE 6437 and FCE 7062. A. Marone 's postdoctoral program was funded by the Marie Curie Intra European Fellowship WASTE2BIOHY (FP7-MC-IEF-326974) under the 7th Framework Programme of the European Community. A. Galès was funded by the French ANR Phycover project (ANR-14-CE04-0011).

Conflict of interest

The authors declare no conflict of interest.

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 110:1550–1160

Ács N, Kovács E, Wirth R, Bagi Z, Strang O, Herbel Z, Rákhely G, Kovács KL (2013) Changes in the Archaea microbial community when the biogas fermenters are fed with protein-rich substrates. Bioresour Technol 131:121–127

Ali Shah F, Mahmood Q, Maroof Shah M, Pervez A, Ahmad Asad S (2014) Microbial ecology of anaerobic digesters: The key players of anaerobiosis. Scientific World Journal 183752

Altschul SF, Gish W, Miller W, Myers E, Lipman D (1990) "Basic local alignment search tool". J Mol Biol 215:403–410

Alvarado A, Montañez-Hernández LE, Palacio-Molina SL, Oropeza-Navarro R, Luévanos-Escareño MP, Balagurusamy N (2014) Microbial trophic interactions and *mcrA* gene expression in monitoring of anaerobic digesters. Front Microbiol 12:1–14

Amann R, Fuchs BM, Behrens S (2001) The identification of microorganisms by fluorescence *in situ* hybridisation. Curr Opin Biotechnol 12:231–236

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

Amann RI (1995) *In situ* identification of microorganisms by whole cell hybridization with rRNA-targeted nucleic acid probes In: Akkermans, ADL, van Elsas, JD, de Bruijn, FJ(Eds), Molecular Microbial Ecology Manual Kluwer Academic Publications, London, pp MMEM-336/1–MMEM-336/15

Amann RI, Zarda B, Stahl DA, Schleifer KH (1992) Identification of individual prokaryotic cells by using enzyme-labeled, rRNA-targeted oligonucleotide probes. Appl Environ Microbiol 58:3007–3011

1247 Araujo JC, Mortara R, Campos JR, Vazoller RF (2004) Development and analysis of anaerobic
1248 biofilms onto hydrophobic and hydrophilic surfaces. Environ Technol 25:809–817

1249 Araujo JC, Brucha G, Campos JR, Vazoller RF (2000) Monitoring the development of
1250 anaerobic biofilms using fluorescent *in situ* hybridization and confocal laser scanning
1251 microscopy. Wat Sci Technol 41:69–77

1252 Araujo JC, Schneider RP (2008) DGGE with genomic DNA: Suitable for detection of
1253 numerically important organisms but not for identification of the most abundant organisms.
1254 Water Res 42:5002–5010

1255 Ariesyady HD, Ito T, Yoshiguchi K, Okabe S (2007) Phylogenetic and functional diversity of
1256 propionate-oxidizing bacteria in an anaerobic digester sludge. Appl Microbiol Biotechnol
1257 75:673–683

1258 Arooj MF, Han SK, Kim SK, Kim DH, Shin HS (2007) Sludge characteristics in anaerobic SBR
1259 system producing hydrogen gas. Water Res 41:1177–1184

1260 Batstone DJ, Keller J, Steyer JP (2006) A review of ADM1 extensions, applications and
1261 analysis. Wat Sci Technol 54:1–10

1262 Batstone DJ and Virdis B (2014) The role of anaerobic digestion in the emerging energy
1263 economy. Curr Opin Biotech 27:142–149

1264 Bauer C, Korthals M, Gronauer A, Lebuhn M (2008) Methanogens in biogas production from
1265 renewable resources - a novel molecular population analysis approach. Water Sci Technol
1266 58:1433–1439

1267 Bekele AZ, Koike S, Kobayashi Y (2011) Phylogenetic diversity and dietary association of
1268 rumen Treponema revealed using group-specific 16S rRNA gene-based analysis. FEMS
1269 Microbiol Lett 316:51–60

1270 Bergmann I, Mundt K, Sontag M, Baumstark I, Nettmann E, Klocke M (2010) Influence of
1271 DNA isolation on Q-PCR-based quantification of methanogenic archaea in biogas fermenters.
1272 Syst Appl Microbiol 33:78–84

1273 Blasco L, Kahala M, Tampio E, Ervasti S, Paavola T, Rintala J, Joutsjoki V (2014) Dynamics of
1274 microbial communities in untreated and autoclaved food waste anaerobic digesters. Anaerobe
1275 29:3–9

1276 Boon N, De Windt W, Verstraete W, Top EM (2002) Evaluation of nested PCR-DGGE
1277 (denaturing gradient gel electrophoresis) with group specific 16S rRNA primers for the analysis
1278 of bacterial communities from different wastewater treatment plants. FEMS Microbiol Ecol
1279 39:101–112

1280 Borzacconi L, Lopez I, Passeggi M, Etchebehere C, Barcia R (2008) Sludge deterioration in a
1281 full scale UASB reactor after a pH drop working under low loading conditions. Wat Sci
1282 Technol 57:797–802

1283 Brankatschk R, Töwe S, Kleineidam K, Schloter M, Zeyer J (2011) Abundances and potential
1284 activities of nitrogen cycling microbial communities along a chronosequence of a glacier
1285 forefield. ISME J 5:1025–1037

1286 Brosius J, Dull TJ, Sleeter DD, Noller HF (1981) Gene organization and primary structure of a
1287 ribosomal RNA operon from *Escherichia coli*. J Mol Biol 148:107–127

Buttigieg PL, Ramette A (2014) A Guide to Statistical Analysis in Microbial Ecology: a community-focused, living review of multivariate data analyses. FEMS Microbiol Ecol 90: 543–550

Calli B, Mertoglu B, Roest K, Inanc B (2006) Comparison of long term performances and final microbial compositions of anaerobic reactor streating landfill leachate. Bioresour Technol 97:641–647

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Gonzalez Peña A, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336

Carballa M, Smits M, Etchebehere C, Boon N, Verstraete W (2011) Correlations between molecular and operational parameters in continuous lab-scale anaerobic reactors. Appl Microbiol Biotechnol 89:303–14

Cardenas E, Tiedje J (2008) New tools for discovering and characterizing microbial diversity. Curr Opin Biotechnol 19:544–549

Castelló E, García y Santos C, Iglesias T, Paolino G, Wenzel J, Borzacconi L, Etchebehere C (2009) Feasibility of biohydrogen production from cheese whey using a UASB reactor: Links between microbial community and reactor performance. Int J Hydrogen Energy 34:5674–5682

Castelló E, Perna V, Wenzel J, Borzacconi L, Etchebehere C (2011) Microbial community composition and reactor performance during hydrogen production in a UASB reactor fed with raw cheese whey inoculated with compost. Wat Sci Technol 64:2265–2273

Čater M, Fanedl L, Logar RM (2013) Microbial community analyses in biogas reactors by molecular methods. Acta Chim Slov 60:243–255

Chang JJ, Wu JH, Wen FS, Hung KY, Chen YT, Hsiao CL, Lin CY, Huang CC (2008) Molecular monitoring of microbes in a continuous hydrogen-producing system with different hydraulic retention time. Int J Hydrogen Energy 33:1579–1585

Chapleur O, Wu TD, Guerquin-Kern JL, Mazeas L, Bouchez T (2013) SIMSISH technique does not alter the apparent isotopic composition of bacterial cells. Plos One 8:e77522

Chen Z, Wang Y, Li K, Zhou H (2014) Effects of increasing organic loading rate on performance and microbial community shift of an up-flow anaerobic sludge blanket reactor treating diluted pharmaceutical wastewater. Journ Biosci Bioeng 118:284–288

Chouari R, Le Paslier D, Dauga C, Daegelen P, Weissenbach J, Sghir A (2005) Novel major bacterial candidate division within a municipal anaerobic sludge digester. Appl Environ Microbiol 71:2145–2153

Ciotola RJ, Martin JF, Castańo JM, Lee J, Michel F (2013) Microbial community response to seasonal temperature variation in a small-scale anaerobic digester. Energies 6:5182–5199

Cole J R, Wang Q, Fish J A, Chai B, McGarrell D M, Sun Y, Brown C T, Porras-Alfaro A, Kuske C R, Tiedje J M (2014) Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res 42:D633–D642

Collins G, Woods A, McHugh S, Carton MW, O'Flaherty V (2003) Microbial community structure and methanogenic activity during start-up of psychrophilic anaerobic digesters treating synthetic industrial wastewaters. FEMS Microbiol Ecol 46:159–70

Connaughton S, Collins G, O´Flaherty V (2006) Development of microbial community structure and activity in a high-rate anaerobic bioreactor at 18°C. Water Res 40:1009–1017

Crocetti G, Murto M, Björnsson L (2006) An update and optimization of oligonucleotide probes targeting methanogenic *Archaea* for use in fluorescence *in situ* hybridization (FISH). J Microbiol Methods 65:194–201

Dar SA, Kuenen JG, Muyzer G (2005) Nested PCR-denaturing gradient gel electrophoresis approach to determine the diversity of sulfate-reducing bacteria in complex microbial communities. Appl Environ Microbiol 71:2325–2330

Davenport RJ, Curtis TP, Goodfellow M, Stainsby FM, Bingley M (2000) Quantitative use of fluorescent *in situ* hybridization to examine relationships between mycolic acid-containing actinomycetes and foaming in activated sludge plants. Appl Environ Microbiol 66:1158–1166

de Gregoris TB, Aldred N, Clare AS, Burgess JG (2011) Improvement of phylum- and class-specific primers for real-time PCR quantification of bacterial taxa. J Microbiol Meth 86:351–356

Delbès C, Moletta R, Godon JJ (2000) Monitoring of activity dynamics of an anaerobic digester bacterial community using 16S rRNA polymerase chain reaction-single-strand conformation polymorphism analysis. Environ Microbiol 2:506–515

Delbès C, Leclerc M, Zumstein E, Godon JJ, Moletta R (2001) A molecular method to study population and activity dynamics in anaerobic digesters. Wat Sci Technol 43:51–57

Díaz EE, Stams AJA, Amils R, Sanz JL (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

Díaz I, Pérez SI, Ferrero EM, Fdz-Polanco M (2011) Effect of oxygen dosing point and mixing on the microaerobic removal of hydrogen sulphide in sludge digesters. Bioresour Technol 102:3768–3775

Egert M, Friedrich MW (2003) Formation of pseudo-Terminal Restriction Fragments, a PCR-related bias affecting Terminal Restriction Fragment Length Polymorphism analysis of microbial community structure. Appl Environ Microbiol 69:2555–2562

Egli K, Bosshard F, Werlen C, Lais P, Siegrist H, Zehnder AJB, Van der Meer JR (2003) Microbial composition and structure of a rotating biological contactor biofilm treating ammonium-rich wastewater without organic carbon. Microb Ecol 45:419–432

Feng XM, Karlsson A, Svensson BH, Bertilsson S (2010) Impact of trace element addition on biogas production from food industrial waste-linking process to microbial communities. FEMS Microbiol Ecol 74:226–40

Fernández A, Huang S, Seston S, Xing J, Hickey R, Criddle C, Tiedje J (1999) How stable is stable? Function versus community composition Appl Environ Microbiol 65:3697–3704

Fernández N, Sierra-Alvarez R, Field JA, Amils R, Sanz JL (2008) Microbial community dynamics in a chemolithotrophic denitrification reactor inoculated with methanogenic granular sludge. Chemosphere 70:462–474

1370 Ferraz ADN, Wenzel J, Etchebehere C, Zaiat M (2014) Effect of organic loading rate on
1371 hydrogen production from sugarcane vinasse in thermophilic acidogenic packed bed reactors.
1372 Int J Hydrog 39:16852–16862

1373 Ferry JG (1999) Enzymology of one-carbon metabolism in methanogenic pathways. FEMS
1374 Microbiol Rev 23:13–38

1375 Franke-Whittle IH, Goberna M, Pfister V, Insam H (2009) Design and development of the
1376 ANAEROCHIP microarray for investigation of methanogenic communities. J Microbiol Meth
1377 79:279–288

1378 Garcia JL, Patel BKC, Ollivier B (2000) Taxonomic, phylogenetic, and ecological diversity of
1379 methanogenic Archaea. Anaerobe 6:205–226

1380 Ginige MP, Hugenholtz P, Daims H, Wagner M, Keller J, Blackall LL (2004) Use of stable-
1381 isotope probing, full-cycle rRNA analysis, and fluorescence *in situ* hybridization
1382 microautoradiographyto study a methanol-fed denitrifying microbial community. Appl Environ
1383 Microbiol 70:558–596

1384 Goberna M, Gadermaier M, Franke-Whittle IH, García C, Wett B, Insam H (2015) Start-up
1385 strategies in manure-fed biogas reactors: Process parameters and methanogenic communities.
1386 Biomass Bioenerg 75:46–56.

1387 Goberna M, Insam H, Franke-Whittle IH (2009) Effect of biowaste sludge maturation on the
1388 diversity of thermophilic bacteria and archaea in an anaerobic reactor. Appl Environ Microbiol
1389 75:2566–2572

1390 Godon JJ, Zumstein E, Dabert P, Habouzit F, Moletta R (1997). Molecular microbial diversity
1391 of an anaerobic digestor as determined by small-subunit rDNA sequence analysis. Appl Environ
1392 Microbiol 63:2802–2813

1393 Griffinths RI, Whiteley AS, O'donnell AG (2000) Rapid method for coextration of DNA from
1394 natural environments for analysis of ribosomal DNA and rRNA-based microbial community
1395 composition. Appl Environ Microbiol 66:88–91

1396 Grosskopf R, Janssen PH, Liesack W (1998) Diversity and structure of the methanogenic
1397 community in anoxic rice paddy soil microcosms as examined by cultivation and direct 16S
1398 rRNA gene sequence retrieval. Appl Environ Microbiol 64:960–969

1399 Gudnason H, Dufva M, Bang DD, Wolff A (2007) Comparison of multiple DNA dyes
1400 for real-time PCR: effects of dye concentration and sequence composition on DNA
1401 amplification and melting temperature. Nucleic Acids Res 35:e127

1402 Hammer ØH (2001) PAST: paleontological statistics software package for education and data
1403 analysis. Palaeontol Electron 4:1–9

1404 Hanreich A, Schimpf U, Zakrzewski M, Schlüter A, Benndorf D, Heyerd R, Rapp E, Pühler A,
1405 Reichl U, Klocke M (2013) Metagenome and metaproteome analyses of microbial communities
1406 in mesophilic biogas-producing anaerobic batch fermentations indicate concerted plant
1407 carbohydrate degradation. Syst Appl Microbiol 36:330–338

1408 Hattori S, Kamagata Y, Hanada S, Shoun H (2000) *Thermacetogenium phaeum* gen nov, sp
1409 nov, a strictly anaerobic, thermophilic, syntrophic acetate-oxidizing bacterium. Int J Syst Evol
1410 Micr 50:1601–1609

1411 Hatzenpichler R, Scheller S, Tavormina PL, Babin BM, Tirrell DA, Orphan VJ (2014) In situ
1412 visualization of newly synthesized proteins in environmental microbes using amino acid tagging
1413 and click chemistry. Environ microbiol 16:2568–2590

1414 Heid CA, Stevens J, Livak KJ, Williams PM (1996) Real time quantitative PCR.
1415 Genome Res 6:986–994

1416 Higuchi R, Dollinger G, Walsh PS, Griffith R (1992) Simultaneous amplification and
1417 detection of specific DNA sequences. Biotechnology (NY) 10:413–417

1418 Hill TCJ, Walsh KA, Harris JA, Moffett BF (2003) Using ecological diversity measures with
1419 bacterial communities. FEMS Microbiol Ecol 43:1–11

1420 Huang WE, Stoecker K, Griffiths R, Newbold L, Daims H, Whiteley AS, Wagner M (2007)
1421 Raman-FISH: combining stable-isotope Raman spectroscopy and fluorescence in situ
1422 hybridization for the single cell analysis of identity and function. Environ microbiol 9:1878–
1423 1889

1424 Imachi H, Sekiguchi Y, Kamagata Y, Ohashi A, Harada H (2000) Cultivation and *in situ*
1425 detection of a thermophilic bacterium capable of oxidizing propionate in syntrophic association
1426 with hydrogenotrophic methanogens in a thermophilic methanogenic granular sludge. Appl
1427 Environ Microbiol 66:3608– 3615

1428 Ito T, Yoshiguchi K, Ariesyady HD, Okabe S (2011) Identification of a novel acetate-utilizing
1429 bacterium belonging to *Synergistes* group 4 in anaerobic digester sludge. ISME J 5:1844–1856

1430 Ito T, Yoshiguchi K, Ariesyady HD, Okabe S (2012) Identification and quantification of key
1431 microbial trophic groups of methanogenic glucose degradation in an anaerobic digester sludge.
1432 Bioresour Technol 123:599–607

1433 Jaenicke S, Ander C, Bekel T, Bisdorf R, Dröge M, Gartemann K-H, Jünemann S, Kaiser O,
1434 Krause L, Tille F, Zakrzewski M, Pühler A, Schlüter A, Goesmannet A (2011) Comparative and
1435 joint analysis of two metagenomic datasets from a biogas fermenter obtained by 454-
1436 pyrosequencing. Plos One 6:e14519

1437 Jang HM, Kim JH, Ha JH, Park JM (2014) Bacterial and methanogenic archaeal communities
1438 during the single-stage anaerobic digestion of high-strength food wastewater. Bioresour Technol
1439 165:174-182

1440 Jaspers E, Overmann J (2004) Ecological significance of microdiversity: identical 16S rRNA
1441 gene sequences can be found in bacteria with highly divergent genomes and ecophysiologies.
1442 Appl Environ Microbiol 70:4831–4839

1443 Jáuregui-Jáuregui JA, Méndez-Acosta HO, González-Álvarez V, Snell-Castro R, Alcaraz-
1444 González V, Godon JJ (2014) Anaerobic treatment of tequila vinasses under seasonal operating
1445 conditions: start-up, normal operation and restart-up after a long stop and starvation period.
1446 Bioresour Technol 168:33–40

1447 Jetten MSM, Strous M, van de Pas-Schoonen KT, Schalk J, van Dongen UGJM, van de Graaf
1448 AA, Logemann S, Muyzer G, van Loosdrecht MCM, Kuenen JG (1999) The anaerobic
1449 oxidation of ammonium FEMS Microbiol Rev 22:421–437

1450 Kampmann K, Ratering S, Baumann R, Schmidt M, Zerr W, Schnell S (2012)
1451 Hydrogenotrophic methanogens dominate in biogas reactors fed with defined substrates. Syst
1452 Appl Microbiol 35:404–413

1453 Karakashev D, Batstone DJ, Trably E, Angelidaki I (2006) Acetate oxidation is the dominant
1454 methanogenic pathway from acetate in the absence of *Methanosaetaceae*. Appl Environ
1455 Microbiol 72:5138–5141

1456 Kim J, Lim J, Lee C (2013) Quantitative real-time PCR approaches for microbial community
1457 studies in wastewater treatment systems: Applications and considerations. Biotechnol adv
1458 31:1358–1373

1459 Klappenbach JA, Dunbar JM, Schmidt TM (2000) rRNA operon copy number reflects
1460 ecological strategies of bacteria. Appl Environ Microb 66:1328–1333

1461 Kundu K, Bergmann I, Hahnke S, Klocke M, Sharma S, Sreekrishnan TR (2013a) Carbon
1462 source: A strong determinant of microbial community structure and performance of an
1463 anaerobic reactor. J Biotechnol 168:616–624

1464 Kundu K, Sharma S, Sreekrishnan T R (2013b) Changes in microbial communities in a hybrid
1465 anaerobic reactor with organic loading rate and temperature Bioresour Technol 129:538–547

1466 Kundu K, Sharma S, Sreekrishnan TR (2012) Effect of operating temperatures on the microbial
1467 community profiles in a high cell density hybrid anaerobic bioreactor. Bioresour Technol
1468 118:502–511

1469 Lane DJ (1991) 16S/23S rRNA sequencing In: Stackebrandt E, Goodfellow M (Eds) Nucleic
1470 Acid Techniques in Bacterial Systematics, Wiley, Chichester, England, pp 115–175

1471 LaPara TM, Nakatsu CH, Pantea L, Alleman JE (2000) Phylogenetic analysis of bacterial
1472 communities in mesophilic and thermophilic bioreactors treating pharmaceutical wastewater.
1473 Appl Environ Microbiol 66:3951–3959

1474 Leclerc M, Delbes C, Moletta R, Godon JJ (2001) Single strand conformation polymorphism
1475 monitoring of 16S rDNA Archaea during start-up of an anaerobic digester. FEMS Microbiol
1476 Ecol 34:213–220

1477 Leclerc M, Delgènes JP, Godon JJ (2004) Diversity of the archaeal community in 44 anaerobic
1478 digesters as determined by single strand conformation polymorphism analysis and 16S rDNA
1479 sequencing. Environ Microbiol 6:809–819

1480 Lee DH, Zo YG, Kim SJ (1996) Nonradioactive method to study genetic profiles of natural
1481 bacterial communities by PCR-single-strand-conformation polymorphism. Appl Environ
1482 Microbiol 62:3112–20

1483 Lee N, Nielsen PH, Andreasen KH, Juretschko S, Nielsen JL, Schleifer KH, Wagner M (1999)
1484 Combination of fluorescent in situ hybridization and microautoradiography - a new tool for
1485 structure-function analyses in microbial ecology. Appl Environ Microbiol 65:1289–1297

1486 Lee C, Kim J, Hwang K, O'Flaherty V Hwang S (2009) Quantitative analysis of methanogenic
1487 community dynamics in three anaerobic batch digesters treating different wastewaters. Water
1488 Res 43:157–165

1489 Levén L, Eriksson ARB, Schnürer A (2007) Effect of process temperature on bacterial and
1490 archaeal communities in two methanogenic bioreactors treating organic household waste. FEMS
1491 Microbiol Ecol 59:683–693

1492 Li A, Chu Y, Wang X, Ren L, Yu J, Liu X, Yan J, Zhang L, Wu S, Li S. (2013) A
1493 pyrosequencing-based metagenomic study of methane-producing microbial community in solid-
1494 state biogas reactor. Biotechnol Biofuels 6:3

Li T, Mazeas L, Sghir A, Leblon G, Bouchez T (2009) Insights into networks of functional microbes catalysing methanization of cellulose under mesophilic conditions. Environ microbiol 11:889–904

Li T, Wu TD, Mazeas L, Toffin L, Guerquin-Kern JL, Leblon G, Bouchez T (2008) Simultaneous analysis of microbial identity and function using NanoSIMS. Environ microbiol 10:580–588

Li M, Gu JD (2011) Advances in methods for detection of anaerobic ammonium oxidizing (anammox) bacteria. Appl Microbiol Biotechnol 90:1241–1252

Limam RD, Chouari R, Mazéas L, Wu T, Li T, Grossin-Debattista J, Guerquin-Kern JL, Saidi M, Landouli A, Sghir A, Bouchez T (2014) Members of the uncultured bacterial candidate division WWE1 are implicated in anaerobic digestion of cellulose. Microbiology Open 3:157–167

Liu WT, Chan OC, Fang HHP (2002) Microbial community dynamics during start-up of acidogenic anaerobic reactors. Water Res 36:3203–3210

Liu W, Marsh TL, Cheng H, Forney LJ (1997) Characterization of microbial diversity by determining Terminal Restriction Fragment Length Polymorphisms of genes encoding 16S rRNA. Appl Environ Microbiol 63:4516–4522

Loisel P, Harmand J, Zemb O, Latrille E, Lobry C, Delgenès JP, Godon JJ (2006) Denaturing gradient electrophoresis (DGE) and single-strand conformation polymorphism (SSCP) molecular fingerprintings revisited by simulation and used as a tool to measure microbial diversity Environ Microbiol 8:720–731

Lü F, Bize A, Guillot A, Monnet V, Madigou C, Chapleur O, Mazéas L, He P, Bouchez T (2014) Metaproteomics of cellulose methanisation under thermophilic conditions reveals a surprisingly high proteolytic activity. ISME J 8:88–102

Lueders T, Chin KJ, Conrad R, Friedrich M (2001) Molecular analyses of methyl-coenzyme M reductase alpha-subunit (*mcrA*) genes in rice field soil and enrichment cultures reveal the methanogenic phenotype of a novel archaeal lineage Environ Microbiol 3:194–204

Luo G, Angelidaki I (2012) Integrated biogas upgrading and hydrogen utilization in an anaerobic reactor containing enriched hydrogenotrophic methanogenic culture. Biotechnol Bioeng 109:2729–2736

Luton PE, Wayne JM, Sharp RJ, Riley PW (2002) The *mcrA* gene as an alternative to 16S rRNA in the phylogenetic analysis of methanogen populations in landfill. Microbiology 148:3521–3530

Ma J, Zhao B, Frear C, Zhao Q, Yu L, Li X, Chen S (2013) *Methanosarcina* domination in anaerobic sequencing batch reactor at short hydraulic retention time. Bioresour Technol 137:41–50

Manz W, Amann R, Ludwig W, Wagner M, Schleifer KH (1992) Phylogenetic oligodeoxynucleotide probes for the major subclasses of proteobacteria: problems and solutions. Syst. Appl. Microbiol. 15:593– 600

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza

JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376–380

Martin-Laurent F, Philippot L, Hallet S, Chaussod R, Germon JC, Soulas G, Catroux G (2001) DNA extraction from soils: old bias for new microbial diversity analysis methods. Appl Environ Microbiol 67:2354–2359

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

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

McInerney MJ, Struchtemeyer CG, Sieber J, Mouttaki H, Stams AJM, Schink B, Rholin L, Gunsalus RP (2008) Physiology, ecology, phylogeny, and genomics of microorganisms capable of syntrophic metabolism Ann NY Acad Sci 1125:58–72

McKeown R, Scully C, Enright AM, Chinalia F, 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

Meyer M, Stenzel U, Myles S, Prufer K, Hofreiter M (2007) Targeted high-throughput sequencing of tagged nucleic acid samples. Nucleic Acids Res 35:e97

Michelland RJ, Dejean S, Combes S, Fortun-Lamothe L, Cauquil L (2009) StatFingerprints: a friendly graphical interface program for processing and analysis of microbial fingerprint profiles. Mol Ecol Resour 9:1359–1363

Mincer TJ, Church MJ, Taylor LT, Preston C, Karl DM, DeLong EF (2007) Quantitative distribution of presumptive archaeal and bacterial nitrifiers in Monterey Bay and the North Pacific Subtropical Gyre. Environ Microbiol 9:1162–1175

Miura Y, Hiraiwa MN, Ito T, Itonaga T, Watanabe Y, Okabe S (2007) Bacterial community structures in MBRs treating municipal wastewater: relationship between community stability and reactor performance. Water Res 41:627–637

Mnif S, Zayen A, Karray F, Bru-Adan V, Loukil S, Godon JJ, Chamkha M, Sayadi S (2012) Microbial population changes in anaerobic membrane bioreactor treating landfill leachate monitored by single-strand conformation polymorphism analysis of 16S rDNA gene fragments. Int Biodet Biodeg 73:50–59

Morgenroth E, Obermayer A, Arnold E, Bruhl A, Wagner M, Wilderer PA (2000) Effect of long-term idle periods on the performance of sequencing batch reactors. Wat Sci Tech 41:105-113

Morris R, Schauer-Gimenez A, Bhattad U, Kearney C, Struble C A, Zitomer D, Maki JS (2014) Methyl coenzyme M reductase (mcrA) gene abundance correlates with activity measurements of methanogenic H_2/CO_2 -enriched anaerobic biomass: *mcrA* genes and transcripts. Microbial Biotechnol 7:77–84

1579 Munk B, Bauer C, Gronauer A, Lebuhn M (2010) Population dynamics of methanogens during
1580 acidification of biogas fermenters fed with maize silage. Eng Life Sci 10:496–508

1581 Muyzer G, de Waal EC, Uitterlinden AG (1993) Profiling of complex microbial populations by
1582 denaturating gradient gel electrophoresis analysis of polymerase chain reaction amplified genes
1583 coding for 16S rRNA. Appl Environ Microbiol 59:695–700

1584 Narihiro T and Sekiguchi Y (2011) Oligonucleotide primers, probes and molecular methods for
1585 the environmental monitoring of methanogenic archaea: probes for methanogens. Microbial
1586 Biotechnol 4:585–602

1587 Nelson MC, Morrison M, Yu Z (2011) A meta-analysis of the microbial diversity observed in
1588 anaerobic digesters. Bioresource Technol 102:3730–3739

1589 Ni S-Q, Zhang J (2013) Anaerobic ammonium oxidation: from laboratory to full-scale
1590 application Bio Med Res Int 469360

1591 Nobu MK, Narihiro T, Rinke C, Kamagata Y, Tringe SG, Woyke T, Liu WT (2015) Microbial
1592 dark matter ecogenomics reveals complex synergistic networks in a methanogenic bioreactor.
1593 ISME J doi: 10.1038/ismej.2014.256

1594 Nocker A, Burr M, Camper AK (2007) Genotypic microbial community profiling: a critical
1595 technical review. Microb Ecol 54:276–289

1596 Noyola A, Padilla-Rivera A, Morgan-Sagastume J M, Güereca L P, Hernández-Padilla F (2012)
1597 Typology of municipal wastewater treatment technologies in latin america Clean, Soil Air
1598 Water 40:926-932

1599 Nübel U, Engelen B, Felske A, Snaidr J, Wieshuber A, Amann RI, Ludwig W, Backhaus H
1600 (1996) Sequence heterogeneities of genes encoding 16S rRNAs in *Paenibacillus polymyxa*
1601 detected by temperature gradient gel electrophoresis. J Bacteriol 178:5636–5643

1602 Okada DY, Delforno TP, Etchebehere C, Varesche MBA (2014) Evaluation of the microbial
1603 community of upflow anaerobic sludge blanket reactors used for the removal and degradation of
1604 linear alkylbenzene sulfonate by pyrosequencing Int Biodeter Biodeg 96:63–70

1605 Osborn AM, Moore ERB, Timmis KN (2000) An evaluation of terminal-restriction fragment
1606 length polymorphism (T-RFLP) analysis for the study of microbial community structure and
1607 dynamics. Environ Microbiol 2:39–50

1608 Osborne CA, Rees GN, Bernstein Y, Janssen PH (2006) New threshold and confidence
1609 estimates for Terminal Restriction Fragment Length Polymorphism analysis of complex
1610 bacterial communities. Appl Environ Microbiol 72:1270–1278

1611 Padmasiri SI, Zhang J, Fitch M, Norddahl B, Morgenroth E, Raskin L (2007) Methanogenic
1612 population dynamics and performance of an anaerobic membrane bioreactor (AnMBR) treating
1613 swine manure under high shear conditions. Water Res 41:134–44

1614 Parameswaran P, Jalili R, Tao L, Shokralla S, Gharizadeh B, Ronaghi M, Fire AZ (2007) A
1615 pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample
1616 multiplexing. Nucleic Acids Res 35:e130

1617 Patel J B (2001) 16S rRNA gene sequencing for bacterial pathogen identification in the clinical
1618 laboratory. Mol Diagn 6:313–321

Pereira AD, Leal CD, Dias MF, Etchebehere C, Chernicharo CA, de Araújo JC (2014) Effect of phenol on the nitrogen removal performance and microbial community structure and composition of an anammox reactor. Bioresour Technol 166:103–111

Pereira JO, MacConell EFA, Silva SQ, Chernicharo CAL (2012) Granular biomass selection in a double-stage biogas collection UASB reactor: effects on SMA, abundance and diversity of the methanogenic population. Wat Sci Technol 66:2570–2577

Perna V, Castello E, Wenzel J, Zampol C, Lima DMF, Borzacconi L, Varesche MB, Zaiat M, Etchebehere C (2013) Hydrogen production in an upflow anaerobic packed bed reactor used to treat cheese whey. Int J Hydrogen Energ 38:54–62

Pernthaler A, Pernthaler J, Amann R (2002) Fluorescence *in situ* hybridization and catalyzed reporter deposition for the identification of marine bacteria. Appl Environ Microbiol 68:3094–3101

Pervin HM, Batstone DJ, Bond PL (2013) Previously unclassified bacteria dominate during thermophilic and mesophilic anaerobic pre-treatment of primary sludge. Syst Appl Microbiol 36:281–290

Power M, Araújo JC, Meer JRVD, Harms H, Wanner O (1999) Monitoring sulfate-reducing bacteria in heterotrophic biofilms. Wat Sci Technol 39:49–56

Prakash O, Pandey PK, Kulkarni GJ, Mahale KN, Shouche YS (2014) Technicalities and glitches of Terminal Restriction Fragment Length Polymorphism (T-RFLP). Indian J Microbiol 54:255–261

Pycke BFG, Etchebehere C, Van De Caveye P, Negroni A, Verstraete W, Boon N (2011) A time-course analysis of four full-scale anaerobic digesters in relation to the dynamics of change of their microbial communities. Wat Sci Technol 63:769–775

Quéméneur M, Hamelin J, Latrille E, Steyer JP, Trably E (2011a) Functional versus phylogenetic fingerprint analyses for monitoring hydrogen-producing bacterial populations in dark fermentation cultures. Int J Hydrogen Energy 36:3870–3879

Quéméneur M, Hamelin J, Benomar S, Guidici-Orticoni MT, Latrille E, Steyer J-P, Trably E (2011b) Changes in hydrogenase genetic diversity and proteomic patterns in mixed-culture dark fermentation of mono-, di- and tri-saccharides. Int J Hydrogen Energy 36:11654–11665

Quéméneur M, Hamelin J, Latrille E, Steyer JP, Trably E (2010) Development and application of a functional CE-SSCP fingerprinting method based on [Fe–Fe]-hydrogenase genes for monitoring hydrogen-producing *Clostridium* in mixed cultures. Int J Hydrogen Energy 35:13158–13167

Radajewski S, Ineson P, Parekh NR, Murrell JC (2000) Stable-isotope probing as a tool in microbial ecology. Nature 403:646–649

Rademacher A, Zakrzewski M, Schlüter A, Schonberg M, Szczepanowski R, Goesmann A, Pühler A, Klocke M (2012) Characterization of microbial biofilms in a thermophilic biogas system by high-throughput metagenome sequencing. FEMS Microbiol Ecol 79:785–799

Rafrafi Y, Trably E, Hamelin J, Latrille E, Meynial-Salles I, Benomar S, Giudici-Orticoni MT, Steyer JP (2013) Sub-dominant bacteria as keystone species in microbial communities producing bio-hydrogen. Int J Hydrogen Energy 38:4975–4985

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

Ramos I, Pérez R, Fdz-Polanco M (2013) Microaerobic desulphurisation unit: A new biological system for the removal of H_2S from biogas. Biores Technol 142:633–640

Ramos I, Pérez R, Fdz-Polanco M (2014) The headspace of microaerobic reactors: Sulphide-oxidising population and the impact of cleaning on the efficiency of biogas desulphurisation. Biores Technol 158:63–73

Raskin L, Poulsen LK, Noguera DR, Rittmann BE, Stahl DA (1994) Quantification of methanogenic groups in anaerobic biological reactors by oligonucleotide probe hybridization. Appl Environ Microbiol 60:1241–1248

Rastogi G, Sani RK (2011) Molecular techniques to assess microbial community structure, function, and dynamics in the environment In: Ahmad I, Ahmad F, Pichtel J (Eds), Microbes and Microbial Technology: Agricultural and Environmental Applications, pp 29–57

Reeve JN, Nölling J, Morgan RM, Smith DR (1997) Methanogenesis: genes, genomes, and who's on first? J Bacteriol 179:5975–5986

Ren N (2013) AD in China: From technical innovation to full-scale implementation In: Proceedings of World Conference on Anaerobic Digestion, June 25 to June 29, 2013, Santiago de Compostela, Spain

Rincon B, Raposo F, Borja R, Gonzalez JM, Portillo MC Saiz-Jiménez C (2006) Performance and microbial communities of a continuous stirred tank anaerobic reactor treating two-phases olive mill solid wastes at low organic loading rates. J Biotechnol 121:534–543

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

Rocheleau S, Greer CW, Lawrence JR, Cantin C, Laramee L, Guiot SR (1999) Differentiation of *Methanosaeta* concilii and *Methanosarcina barkeri* in anaerobic mesophilic granular sludge by fluorescent *in situ* hybridization and confocal scanning laser microscopy. Appl Environ Microbiol 65:2222–2229

Roest K, Heilig HG, Smidt H, de Vos WM, Stams AJM, Akkermans ADL (2005) Community analysis of a full-scale anaerobic bioreactor treating paper mill wastewater. Syst Appl Microbiol 28:175–185

Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Leamon JH, Johnson K, Milgrew MJ, Edwards M, Hoon J, Simons JF, Marran D, Myers JW, Davidson JF, Branting A, Nobile JR, Puc BP, Light D, Clark TA, Huber M, Branciforte JT, Stoner IB, Cawley SE, Lyons M, Fu Y, Homer N, Sedova M, Miao X, Reed B, Sabina J, Feierstein E, Schorn M, Alanjary M, Dimalanta E, Dressman D, Kasinskas R, Sokolsky T, Fidanza JA, Namsaraev E, McKernan KJ, Williams A, Roth GT, Bustillo J (2011) An integrated semiconductor device enabling non-optical genome sequencing. Nature 475:348–352

Rousseau R, Van Hecke P, Nijssen D, Bogaert J (1999) The relationship between diversity profiles, evenness and species richness based on partial ordering. Environ Ecol Stat 6:211–223

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF

(2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537–7541

Schlüter A, Bekel T, Díaz NN, Dondrup M, Eichenlaub R, Gartemann KH, Krahn I, Krause L, Krömeke H, Kruse O, Mussgnug JH, Neuweger H, Niehaus K, Pühler A, Runte KJ, Szczepanowski R, Tauch A, Tilker A, Viehöver P, Goesmann A (2008) The metagenome of a biogas-producing microbial community of a production-scale biogas plant fermenter analysed by the 454-pyrosequencing technology. J Biotechnol 136:77–90

Schmalenberger A, Tebbe CC (2002) Bacterial diversity in maize rhizospheres: conclusions on the use of genetic profiles based on PCR-amplified partial small subunit rRNA genes in ecological studies. Mol Ecol 12:251–262

Schütte UME, Abdo Z, Bent SJ, Shyu C, Williams CJ, Pierson JD, Forney LJ (2008) Advances in the use of Terminal Restriction Fragment Length Polymorphism (T-RFLP) analysis of 16S rRNA genes to characterize microbial communities. Appl Microbiol Biotechnol 80:365–80

Sekiguchi H, Tomioka N, Nakahara T, Uchiyama H (2001) A single band does not always represent single bacterial strains in denaturating gradient gel electrophoresis analysis. Biotechnol Lett 23:1205–1208

Sekiguchi Y, Takahashi H, Kamagata Y, Ohashi A, Harada H (2001) In situ detection, isolation, physiological properties of a thin filamentous microorganism abundant in methanogenic granular sludges: a novel isolate affiliated with a clone cluster, the green non-sulfur bacteria, subdivision I. Appl Environ Microbiol 67:5740–5749

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

Sekiguchi Y, Kagamata Y, Syutsubo K, Ohashi A, Harada H (1998) Phylogenetic diversity of mesophilic and thermophilic granular sludges determined by 16S rRNA gene analysis. Microbiology 144:2655–2665

Shendure J, Ji H (2008) Next-generation DNA sequencing. Nat Biotechnol 26:1135–1145

Shi J, Zhongjiang W, Stiverson J, Zhongtang Y, Li Y (2013) Reactor performance and microbial community dynamics during solidstate anaerobic digestion of corn stover at mesophilic and thermophilic conditions. Bioresour Technol 136:574–581

Shokralla S, Spall JL, Gibson JF, Hajibabaei M (2012) Next-generation sequencing technologies for environmental DNA research. Mol Ecol 21:1794–1805

Smith KS, Ingram-Smith C (2007) Methanosaeta, the forgotten methanogen? Trends Microbiol 15:150–155

Song M, Shin SG, Hwang S (2010) Methanogenic population dynamics assessed by real-time quantitative PCR in sludge granule in upflow anaerobic sludge blanket treating swine wastewater. Bioresource Technol 101:23–28

Speksnijder AGCL, Kowalchuk GA, De Jong S, Kline E, Stephen JR, Laanbroek HJ (2001) Microvariation artifacts introduced by PCR and cloning of closely related 16S rRNA gene sequences. Appl Environ Microbiol 67:469–472

Stahl DA, Amann R (1991) Development and application of nucleic acid probes in bacterial systematics In: Stackebrandt, E, Goodfellow, M (Eds) Nucleic acid techniques in bacterial systematics, Wiley, Chichester, England, pp 205–248

Stams AJM, Plugge CM (2009) Electron transfer in syntrophic communities of anaerobic bacteria and archaea Nat Rev Microbiol 7:568–577

Stams AJM, Sousa DZ, Kleerebezem R, Plugge CM (2012) Role of syntrophic microbial communities in high-rate methanogenic bioreactors. Wat Sci Technol 66:352–362

Steinberg LM, Regan JM (2009) *mcrA*-targeted real-time quantitative PCR method to examine methanogen communities. Appl Environ Microbiol 75:4435–4442

Su C, Lei L, Duan Y, Zhang K-Q, Yang J (2012) Culture-independent methods for studying environmental microorganisms: methods, application, and perspective. Appl Microbiol Biotechnol 93:993–1003

Sun DL, Jiang X, Wu QL, Zhou NY (2013) Intragenomic heterogeneity of 16S rRNA genes causes overestimation of prokaryotic diversity. Appl Environ Microbiol 79:5962–5969

Sundberg C, Al-Soud WA, Larsson M, Alm E, Yekta SS, Svensson BH, Sørensen SJ, Karlsson A (2013) 454 pyrosequencing analyses of bacterial and archaeal richness in 21 full-scale biogas digesters. FEMS Microbiol Ecol 85:612–626

Suzuki MT, Taylor LT, DeLong EF (2000) Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5′-nuclease assays. Appl Environ Microbiol 66:4605–4614

Takai K, Horikoshi K (2000) Rapid detection and quantification of members of the archaeal community by quantitative PCR using fluorogenic probes. Appl Environ Microbiol 66:5066–5072

Talbot G, Topp E, Palin MF, Massé DI (2008) Evaluation of molecular methods used for establishing the interactions and functions of microorganisms in anaerobic bioreactors. Water Res 42:513–537

Temperton B, Giovannoni SJ (2012) Metagenomics: microbial diversity through a scratched lens. Curr Opin Microbiol 15:605–612

Tian Z, Zhang Y, Li Y, Chi Y, Yang M (2015) Rapid establishment of thermophilic anaerobic microbial community during the one-step startup of thermophilic anaerobic digestion from a mesophilic digester. Water Res 69:9–19

Tolvanen KES, Koskinen PEP, Raussi HM, Ylikoski AI, Hemmilä IA, Santala VP, Karp MT (2008a) Profiling the *hydA* gene and *hydA* gene transcript levels of *Clostridium butyricum* during continuous, mixed-culture hydrogen fermentation. Int J Hydrogen Energy 33:5416–5421

Tolvanen KES, Koskinen PEP, Ylikoski AI, Ollikka PK, Hemmilä IA, Puhakka JA, Karp MT (2008b) Quantitative monitoring of a hydrogen-producing *Clostridium butyricum* strain from a continuous-flow, mixed culture bioreactor employing real-time PCR. Int J Hydrogen Energy 33:542–549

Town JR, Links MG, Fonstad TA, Dumonceaux TJ (2014) Molecular characterization of anaerobic digester microbial communities identifies microorganisms that correlate to reactor performance. Bioresour Technol 151:249-57

Trably E, Batstone DJ, Christensen N, Patureau D, Schmidt JE (2008) Microbial dynamics in anaerobic enrichment cultures degrading di-n-butyl phthalic acid ester. FEMS Microbiol Ecol 66:472–483

Traversi D, Villa S, Acri M, Pietrangeli B, Degan R, Gilli G (2011) The role of different methnogen groups evaluated by real-time qPCR as high-efficiency bioindicators of wet anaerobic co-digestion of organic waste. AMB Express 1:28

Tresse O, Mounien F, Levesque MJ, Guiot S (2005) Comparison of the microbial population dynamics and phylogenetic characterization of a CANOXIS reactor and a UASB reactor degrading trichloroethane. J Appl Microbiol 98:440–449

Tu Q, Yu H, He Z, Deng Y, Wu L, Van Nostrand JD, Zhou A, Voordeckers J, Lee YJ, Qin Y, Hemme CL, Shi Z, Xue K, Yuan T, Wang A, Zhou J (2014) GeoChip 4: A functional gene-array-based high-throughput environmental technology for microbial community analysis. Mol Ecol Resour 14:914-928

Ueno Y, Haruta S, Ishii M, Igarashi Y (2001) Changes in product formation and bacterial community by dilution rate on carbohydrate fermentation by methanogenic microflora in continuous flow stirred tank reactor. Appl Microbiol Biotechnol 57:65–73

Van Elsas JD, Boersma FGH (2011) A review of molecular methods to study the microbiota of soil and the mycosphere. Eur J Soil Biol 47:77–87

Vanwonterghem I, Jensen PD, Ho DP, Batstone DJ, Tyson GW (2014) Linking microbial community structure, interactions and function in anaerobic digesters using new molecular techniques. Curr Opin Biotech 27:55–64

Verstraete W (2007) Microbial ecology and environmental biotechnology. ISME J 1:4–8

Wang MY, Olson BH, Chang JS (2008a) Relationship among growth parameters for *Clostridium butyricum*, *hydA* gene expression, and biohydrogen production in a sucrose-supplemented batch reactor. Appl Microbiol Biotechnol 78:525–532

Wang MY, Tsai YL, Olson BH, Chang JS (2008b) Monitoring dark hydrogen fermentation performance of indigenous *Clostridium butyricum* by hydrogenase gene expression using RT-PCR and qPCR. Int J Hydrogen Energy 33:4730–4738

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. Appl Environ Microbiol 73:5261–5267

Wang Y, Qian PY (2009) Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. PloS One 4:e7401

Werner JJ, Knights D, Garcia ML, Scalfone NB, Smith S, Yarasheski K, Cummings TA, Beers AR, Knight R, Angenent LT (2011) Bacterial community structures are unique and resilient in full-scale bioenergy systems. P Natl Acad Sci USA 108: 4158–4163

Whang LM, Hu TH, Liu PWG, Hung YC, Fukushima T, Wu YJ, Chang SH (2015) Molecular analysis of methanogens involved in methanogenic degradation of tetramethylammonium hydroxide in full-scale bioreactors. Appl Microbiol Biotechnol 99:1485–1497

Wilhartitz I, Mach RL, Teira E, Reinthaler T, Herndl GJ, Farnleitner AH (2007) Prokaryotic community analysis with CARD-FISH in comparison to FISH in ultra-oligotrophic ground- and drinking water. J Appl Microbiol 103:871–881

1825 Wittebolle L, Vervaeren H, Verstraete W, Boon N (2008) Quantifying community dynamics of
1826 nitrifiers in functionally stable reactors. Appl Environ Microbiol 74:286–293

1827 Wittwer CT, Herrmann, MG, Moss, AA, Rasmussen, RP (1997) Continuous fluorescence
1828 monitoring of rapid cycle DNA amplification. Biotechniques 22:130–131

1829 Wong MT, Zhang D, Li J, Hui RK, Tun HM, Brar MS, Park TJ, Chen Y, Leung FC (2013)
1830 Towards a metagenomic understanding on enhanced biomethane production from waste
1831 activated sludge after pH 10 pretreatment Biotechnol Biofuels 6:38

1832 Xu K, Liu H, Chen J (2010) Effect of classic methanogenic inhibitors on the quantity and
1833 diversity of archaeal community and the reductive homoacetogenic activity during the process
1834 of anaerobic sludge digestion. Bioresour Technol 101:2600–2607

1835 Yamada T, Sekiguchi Y, Imachi H, Kamagata Y, Ohashi A, Harada H (2005) Diversity,
1836 localization and physiological properties of filamentous microbes belonging to *Chloroflexi*
1837 subphylum I in mesophilic and thermophilic methanogenic sludge granules. Appl Environ
1838 Microbiol 71:7493–7503

1839 Yamada T, Yamauchi T, Shiraishi K, Hugenholtz P, Ohashi A, Harada H, Kamagata Y,
1840 Nakamura K Sekiguchi Y (2007) Characterization of filamentous bacteria, belonging to
1841 candidate phylum KSB3, that are associated with bulking in methanogenic granular sludges.
1842 ISME J 1:246–255

1843 Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer KH, Whitman WB, Euzéby
1844 J, Amann R, Rosselló-Móra R (2014) Uniting the classification of cultured and uncultured
1845 bacteria and archaea using 16S rRNA gene sequences. Nat Rev Microbiol 12:635–645

1846 Ye NF, Lü F, Shao LM, Godon LM, He PJ (2007) Bacterial community dynamics and product
1847 distribution during pH-adjusted fermentation of vegetable wastes. J Appl Microbiol 103:1055–
1848 65

1849 Yu Y, Lee C, Kim J, Hwang S (2005) Group-specific primer and probe sets to detect
1850 methanogenic communities using quantitative real-time polymerase chain reaction. Biotechnol
1851 Bioeng 89:670–679

1852 Zakrzewski M, Goesmann A, Jaenicke S, Jünemann S, Eikmeyer F, Szczepanowski R, Al-Soud
1853 WA, Sørensen S, Pühler A, Schlüter A (2012) Profiling of the metabolically active community
1854 from a production-scale biogas plant by means of high-throughput metatranscriptome
1855 sequencing. J Biotechnol 158):248–258

1856 Zemb O, Haegeman B, Delgenes JP, Lebaron P, Godon JJ (2007) SAFUM: statistical analysis
1857 of SSCP fingerprints using PCA projections, dendrograms and diversity estimators. Mol Ecol
1858 Notes 7:767–770

1859 Zhang W, Werner JJ, Agler MT, Angenent LT (2014) Substrate type drives variation in reactor
1860 microbiomes of anaerobic digesters. Bioresour Technol 151:397–401.

1861 Zhang C, Yuan Q, Lu Y (2014) Inhibitory effects of ammonia on methanogen *mcr*A transcripts
1862 in anaerobic digester sludge. FEMS Microbiol Ecol 87:368–77

1863 Zhang Y, Aiyuk S, Xu H, Chen G, Verstraete W (2005) Study of microbial community
1864 structures in UASB sludge treating municipal wastewaters by denaturing gradient gel
1865 electrophoresis of 16S rRNA. Sci China C Life Sci 48:128–135

Zhao J, Zuo J, Wang X, Lin J, Yang Y, Zhou J, Chu H, Li P (2014) GeoChip-based analysis of microbial community of a combined nitritation-anammox reactor treating anaerobic digestion supernatant. Water Res 67:345–354

Zheng D, Raskin L (2000) Quantification of *Methanosaeta* species in anaerobic bioreactors using genus- and species- specific hybridization probes. Microb Ecol 39:246–262

Zhu C, Zhang J, Tang Y, Zhengkai X, Song R (2011) Diversity of methanogenic archaea in a biogas reactor fed with swine feces as the mono-substrate by *mcr*A analysis. Microbiol Res 166:27–35

Ziembinska-Buczynska A, Banach A, Bacza T, Pieczykolan M (2014) Diversity and variability of methanogens during shift from mesophilic to thermophilic conditions while biogas production. World J Microbiol Biotechnol 30:3047–3053

Ziganshin AM, Liebetrau J, Proter J, Kleinsteuber S (2013) Microbial community structure and dynamics during anaerobic digestion of various agricultural waste materials. Appl Microbiol Biotechnol 97:5161–5174

Zinder SH (1984) Microbiology of anaerobic conversion of organic wastes to methane: recent developments ASM News 50:294–298

Zumstein E, Moletta R, Godon JJ (2000) Examination of two years of community dynamics in an anaerobic bioreactor using fluorescence polymerase chain reaction (PCR) single-strand conformation polymorphism analysis. Environ Microbiol 2:69–78

Figures revised
Click here to download Figure: Figures reviewed.docx

Figure 1- Schematic view of the anaerobic degradation of organic matter

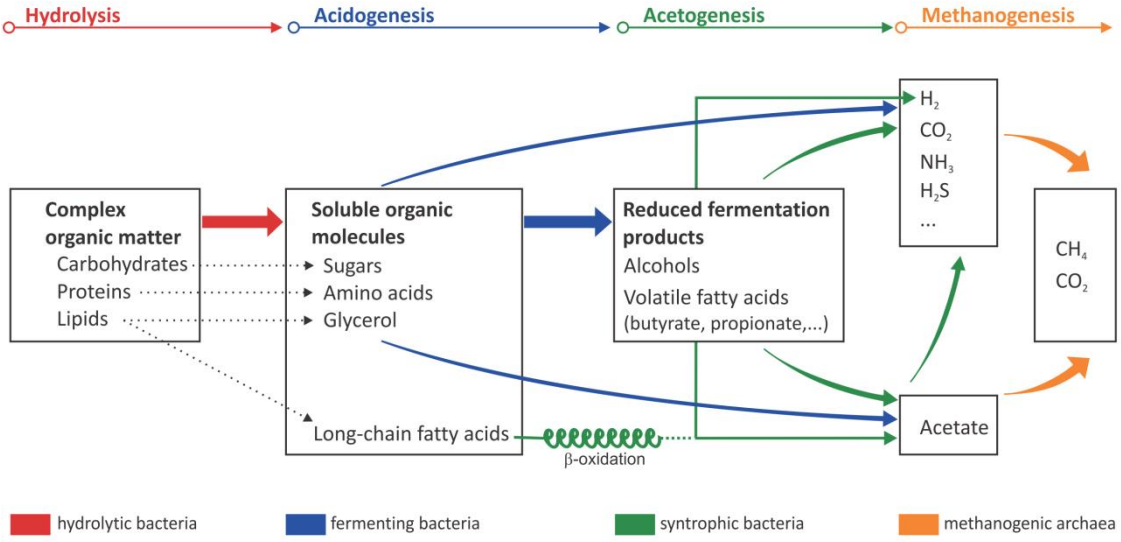
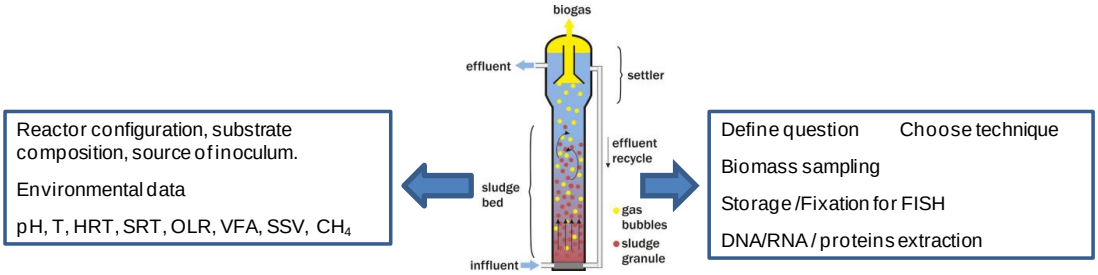


Figure 2- Questions to be formulated and molecular techniques that could be used to answer them, results obtained and statistical analysis.



Question	Who is there?	Changes during operation ?	How many ?	Which is the function ?
Methods	**Cloning sequencing NGS**	**DGGE T-RFLP SSCP**	**Q-PCR FISH**	**Advanced FISH SIP Meta-omics**
Results obtained	Sequences OTUs matrix Taxonomy Phylogenetic tree	Gel/ Chromatogram Band/peak matrix Taxonomy Band/peak	Copy number Biovolume	Identity of active cells Genes Transcripts Proteins Molecules
Statistical analysis	Cluster analysis PCA, CCA Diversity indices Rr, Dy, Fo			

Figure 3 Methods to study microbial community structure and dynamics based on the amplification of 16S rRNA genes or functional genes: DGGE, T-RFLP and SSCP.

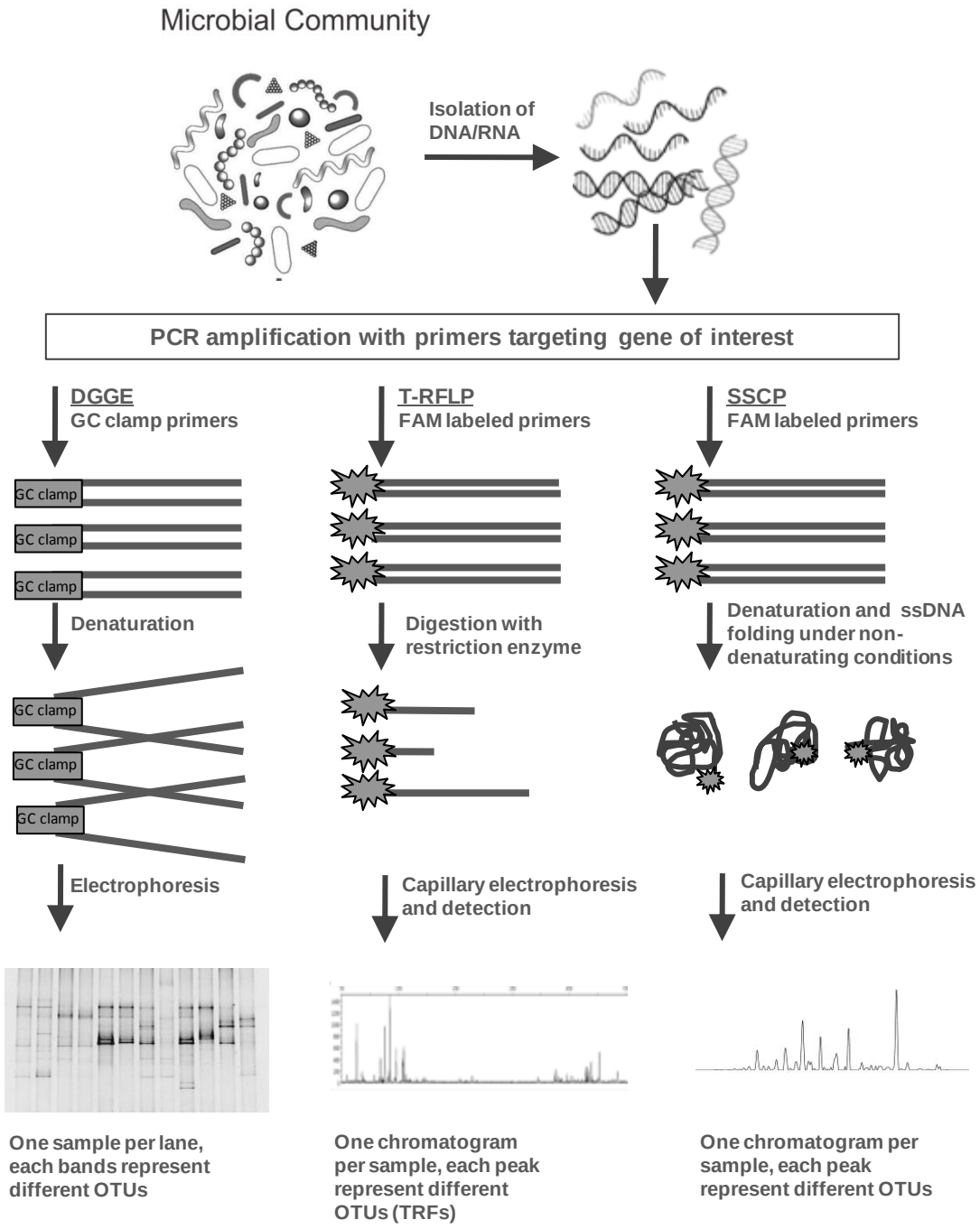


Figure 4. Methods to study microbial diversity and function in complex ecosystems: (a) Stable isotope probing (SIP), and (b)*in situ* microscopic methods coupled to stable/radioactive spectroscopy techniques (microautoradiography coupled to FISH analysis: MAR-FISH; Raman microspectrometry coupled to FISH: Raman-FISH; .nanoscale secondary ion mass spectrometry (NanoSIMS) coupled to FISH (SIMSISH).

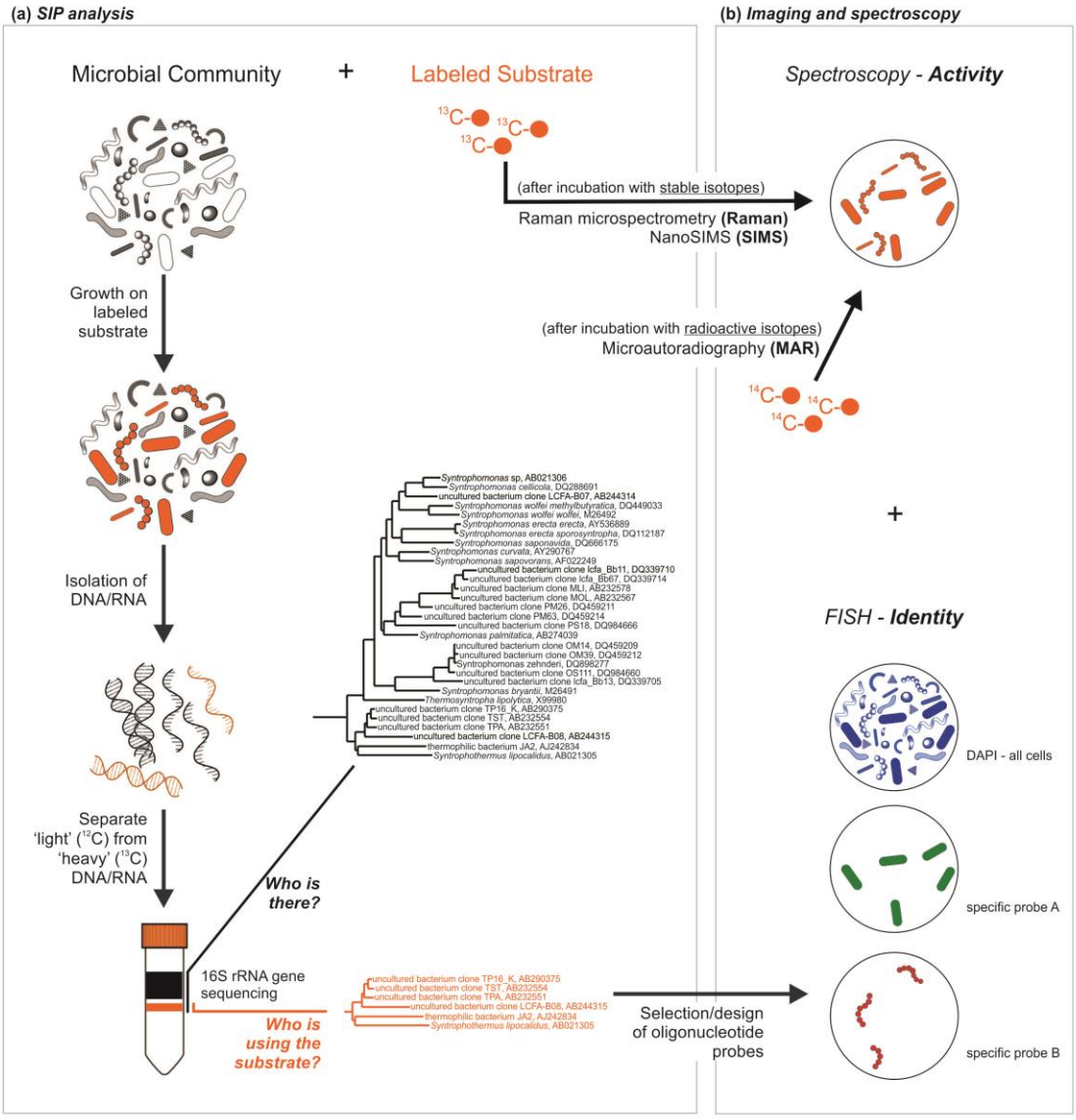


Figure 5. Possibilities for multi-omics analysis of complex microbial communities. Metagenomics analysis provides information on taxonomy (based on retrieved 16S rRNA gene sequences) and on the metabolic genes (potential functional metabolic pathways) present in the metagenome. Metatranscriptomics enables the investigation of the actively transcribed ribosomal and messenger RNA from a community, giving more insight on the active fraction of the community; it can be used to study how communities and gene expression change in response to environmental changes. Metaproteomics is an excellent tool to for studying functionality in an ecosystem because proteins are the direct responsible for cell phenotype and are therefore more informative than the identification of functional genes or their corresponding messenger RNAs. Metabolomics has been recently introduced in systems biology approaches and studies the complete set of metabolites formed by the whole microbial community as a result of its interaction with biotic and abiotic factors of its environment.

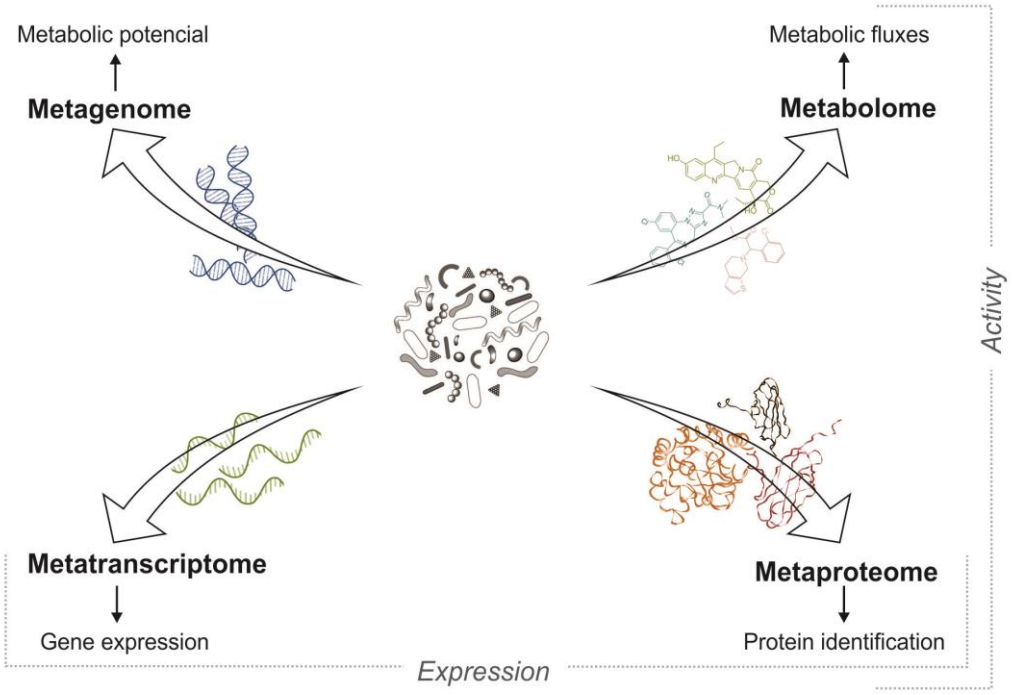


Figure 6. Hydrogenotrophic and acetoclasticmethanogenesis pathways. Genes detected in methagenomic analyses of anaerobic digesters are indicated in blue - data from Rademacher et al.(2012) and Wong et al. (2013); in red, genes that could not be detected in those studies. Dotted lines represent an alternative pathway for the activation of acetate to acetyl-CoA, as suggested by Wong et al. (2013).

MF - methanofuran; MPT - methanopterin; CoM - coenzyme M; CoB - coenzyme B; CoA - coenzyme A; F_{420} - coenzyme F_{420}; Fd - ferredoxin.

fmd - formylmethanofuran dehydrogenase; *ftr* - formylmethanofuran:H_4MPT formyltransferase; *mch* - methenyl-H_4MPT cyclohydrolase; *hmd* - H_2-forming methylene- H_4MPT dehydrogenase; *mer* - F_{420}-dependent methylene-H_4MPT reductase; *mtr* - methyl- H_4MPT:coenzyme M methyltransferase; *mcr* - methyl coenzyme M reductase; *ack* - acetate kinase; *pta* - phosphotransacetylase; *acs*- acetyl-CoA synthetase; *ppa*- inorganic pyrophosphatase; *chd* - carbon monoxide dehydrogenase/acetyl-CoA synthase.

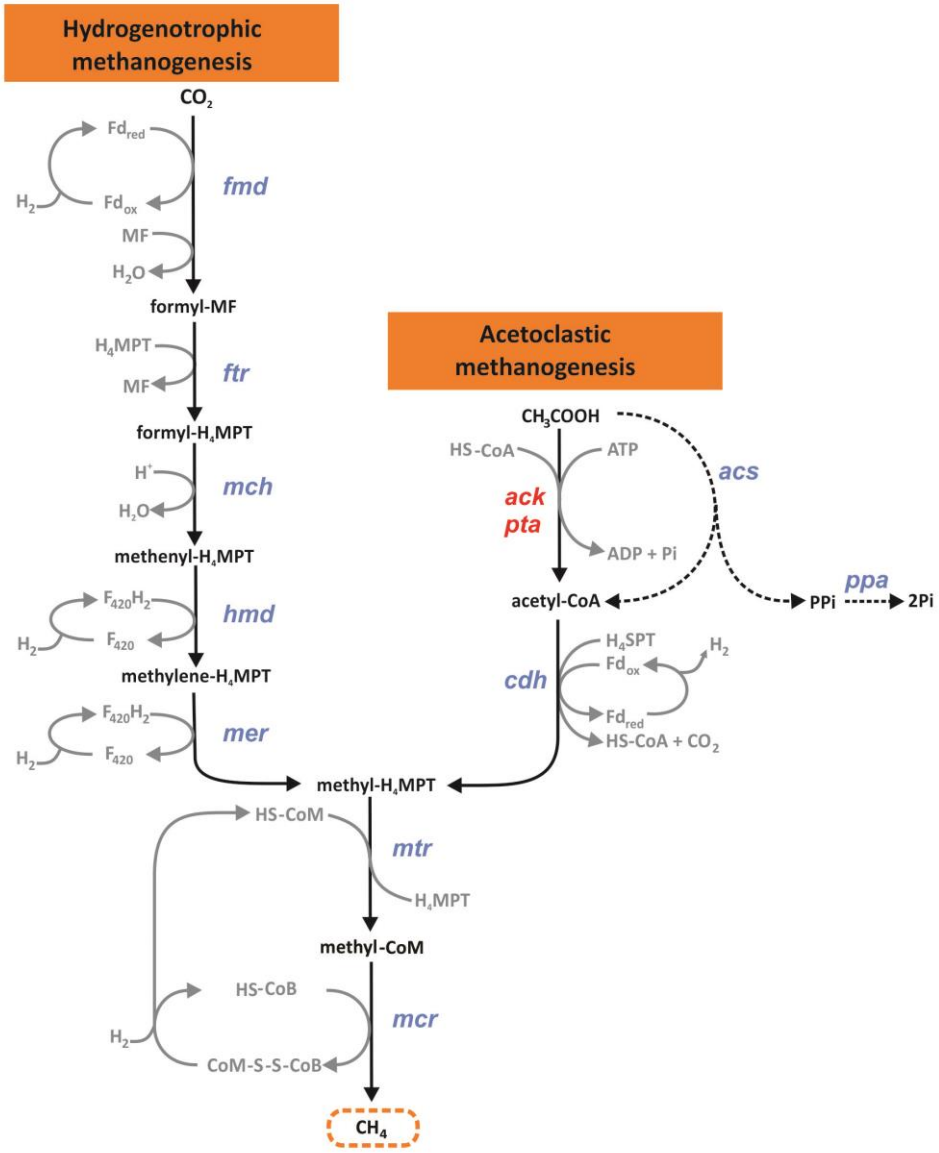


Figure 7- Possible applications of molecular biology techniques results to improve AD processes.

1- Prevent failures- By the regular monitoring of the microbial community of a bioreactor, information about the dynamic and the organization of the community during bad and good performance could be retrieved and used to predict the failures. The knowledge of microorganisms that cause problems (bulking, acidification, etc.) could be also used to prevents their growth..
2- Select inoculum- With the NGS available, it is possible to have a complete view of the microbial composition of a sample in one or two days with low cost this information could help to choose an adequate source of inoculum for full scale reactors avoiding cost of long periods of start up or reinoculation.
3- Data for modeling- Molecular data could be used to design systems of automatic control of reactors and for modeling. The AD model is based in data from kinetic and bulk measurements of the biomass. An effort of conceptualization of the data and integration of the two disciplines should be done.
4- Bioaugmentation- With the knowledge of key microorganisms important for the process it will be possible to design strategies for bioaugmentation and follow the growth of these organisms in the reactor. This could be very useful for the degradation of recalcitrant compounds (like detergents, pharmaceutical products) or to improve the degradation of difficult substrates (as lipids).
5- Design new process-With the discover of new metabolic functions of the microorganisms involved in the AD process it could be think different engineering process to solve some practical issues.

Table 1. Over view of the main bacterial phyla that composed the microbial communities from full scale anaerobic reactors. The values are mean of the values presented in the reports according to the results obtained from 16S rRNA gene sequencing analysis using different methods. Report 1: Nelson et al. (2011), Report 2: Riviere et al. (2009), Report 3: Sunberg et al. (2013).The data from unclassified sequences and some groups with low sequence abundance were not presented.

				Bacterial Phylum (%)								
Report	**Method**	**Reactors analyzed**	**Total sequences within Bacteria**	***Chloroflexi***	***Proteobacteria***	***Firmicutes***	***Bacteroidetes***	***Synergistetes***	***Actinobateria***	***Spirochaetes***	***Verrucomicrobia***	***Termotogales***
1	Sequences retrieved from database	NR*	16,519	23	22	15	15	6	2	1	1	1
2	Cloning/Sanger sequencing	7 UASB reactors treating municipal sewage sludge	7,000	32	18	9	11	6	2	1	1	1
3	454 pyrosequencing	21 CSTR reactors treating solid waste **	164,822									
		7 SS		1.7	5.7	25	15		4.7	4.6	0.9	1.6
		10 CD		0.1	0.4	69	14		1	0.1	0.1	0.3
		4 CDT		0.3	0.2	67	7		0.3	0.1	0.1	19

*NR= not reported.

** The reactors studied were operated in different conditions: 7 reactors were fed with sewage sludge (SS), 10 reactors were fed with compose digestion (CD), and 4 reactors with compose digestion in thermophilic conditions (CDT).

Table 2. Overview of main functional groups detected in anaerobic digestion sludges by using SIP, probing and advanced imaging techniques

Sample origin	Technique	Test substrate	Main bacterial groups detected*	Main archaeal groups detected*	Observations	Reference
Full-scale mesophilic anaerobic municipal sludge digester	MAR-FISH	^{14}C-glucose	Chloroflexi Spirochaetes *Spirochaeta* Candidate phylum TM7		β-Proteobacteria was the most dominant group in butyrate-, propionate- and acetate-degrading communities. MAR-positive β-Proteobacteria in each sample were collected by micromanipulation and their 16S rRNA gene cloned and sequenced. Half of the clones under the three conditions were affiliated with the genus *Variovorax*. *Smithella* and *Syntrophobacter* species play an important role in propionate degradation. Relative numbers of different populations within these genus vary with propionate concentration.	(Ariesyady et al. 2007a; Ariesyady et al. 2007b)
		^{14}C-butyrate	β-Proteobacteria Firmicutes *Syntrophomonas*			
		^{14}C-propionate	β-Proteobacteria δ-Proteobacteria *Smithella* *Syntrophobacter*			
		^{14}C-acetate	β-Proteobacteria γ-Proteobacteria Firmicutes Synergistetes *Synergistes*	Methanosarcinales *Methanosaeta* *Methanosarcina*		
Full-scale mesophilic anaerobic municipal solid waste digester	DNA- SIP	^{13}C-cellulose	Firmicutes *Clostridia, Acetivibrio* *Clostridia,* group 4	Methanomicrobiales *Methanoculleus*	*Acetivibrio* species were detected in aggregates around cellulose fibers using FISH, indicating that these bacteria are major cellulose degraders. Detection of *Methanoculleus* in the 'heavy' DNA may result from the assimilation of marked CO_2 deriving from cellulose fermentation.	(Li et al. 2009)
		^{13}C-glucose	Firmicutes *Clostridia, Clostridium* Bacteroidetes *Porphyromonadaceae* *Rikenellaceae*		Hybridization with specific FISH probes confirmed abundant representation of *Clostridium* and *Porphyromonadaceae* (mainly in the planktonic phase).	
		^{13}C- acetate	γ- Proteobacteria *Pseudomonadaceae*	Methanomicrobiales *Methanoculleus*	FISH confirmed abundant representation of *Pseudomonadaceae*-related bacteria.	

					Syntrophic acetate oxidation between *Pseudomonadaceae*-related bacteria and *Methanoculleus* species was hypothesized.	
Lab-scale mesophilic anaerobic fed-batch reactors fed with powdered whole milk	RNA-SIP MAR-FISH	^{13}C-glucose ^{14}C-glucose	Actinobacteria *Olsenella* *Propionibacterium* Bacteroidetes Chloroflexi Synergistetes *Synergistes*, group 4 δ-Proteobacteria *Syntrophobacter* *Smithella*	Methanosarcinales *Methanosaeta*	Actinobateria, Bacteroidetes and Chloroflexi linked to glucose conversion to propionate and acetate. These labelled intermediates were further assimilated by propionate utilizers (*Syntrophobacter* and *Smithella)* and acetate utilizers (*Synergistes* group 4 and *Methanosaeta*). Actinobacteria was the most abundant MAR-FISH positive microbial group; increase of this group coincided with decreasing glucose concentration. *Smithella* has higher affinity to propionate than *Syntrophobacter* and, therefore, predominated when propionate concentrations were low.	(Ito et al. 2012)
Municipal solid waste digester	DNA-SIP SIMSISH	^{13}C-cellulose	Candidate division WWE1		Implication of candidate division WWE1 in anaerobic digestion of cellulose was confirmed at single-cell level by combined FISH-NanoSIMS analysis. Nevertheless, a major part of ^{13}C-enriched microbes did not exhibit positive hybridization signals with WWE1-specific probe indicating that different community members were using ^{13}C-cellulose.	(Limam et al. 2014)

* For SIP analysis, main microbial groups detected in heavy fractions; for MAR-FISH, main MAR-positive groups are indicated

1 Table 3. Advantages and limits of the molecular methods mostly used in anaerobic digestion; n/a: not appropriate.

Name	**NGS**	**Cloning-sequencing**	**Microarrays**	**Fingerprint (DGGE, T-RFLP, SSCP)**	**qPCR**	**FISH, advanced FISH**	**SIP**	**Transcriptomic**	**Proteomic**
Molecule targeted	Nucleic acids	Nucleic acids	Nucleic acids	Nucleic acids	Nucleic acids	Nucleic acids	Nucleic acids	RNA	Protein
PCR based	Yes	Yes	No	Yes	Yes	No	No	Yes	No
Physiological state of the cells	n/a	n/a	n/a	n/a	n/a	Alive or active	Active	Active	Active
Sampling and storage caution	-	-	-	-	-	Fixation	-	Rapid freezing	Rapid freezing
Specific procedure and equipment	Sequencer Bioinformatic data processing	Sequencer	Pre-designed array	Gel/capillary elecrophoresis	Real-time PCR	Fluorescent/ confocal microscope	density gradient centrifugation	Sequencer Bioinformatic data processing	Mass spectro. Bioinformatic data processing
Technical easiness	+	+	-	+++	+++	++	+	+	+
Data processing requirements	+++	++	++	+	-	-	-	+++	+++
High throughput	+++	-	+	++	+	-	-	++	++
Phylogenetic identification	Yes	Yes	Yes	n/a	Yes	Yes	n/a *	Yes	Yes
Diversity measurement	Yes	Yes	Yes	Yes	n/a	n/a	n/a *	Yes	Yes
Spatial structuration	n/a	n/a	n/a	n/a	n/a	Yes	n/a*	n/a	n/a
Current use in AD	+++	+	-	+++	+++	+	+	+	+
Effort in technical development	+++	-	+	+	++	+	-	++	++

*Yes, if combined with sequencing (for phylogenetic identification and diversity measurement) or imaging (for spatial structuration) – *cf.* figure 4