COMMUNAUTE FRANCAISE DE BELGIQUE

ACADEMIE UNIVERSITAIRE WALLONIE-EUROPE

UNIVERSITE DE LIEGE – GEMBLOUX AGRO-BIO TECH

MANAGEMENT OF MICROBIAL COMMUNITIES ON THE BASIS OF SINGLE CELL TECHNOLOGIES: APPLICATION TO CELLULOSE DEGRADATION AND BIOFILM FORMATION

KINET ROMAIN

Dissertation submitted in fulfilment of the requirements for the degree of Doctor in

Agricultural Sciences and Biological Engineering

Supervisor: Frank Delvigne

June 2016

Kinet Romain (2016): Management of microbial communities on the basis of single cell technologies: application to cellulose degradation and biofilm formation

Thesis summary

Bioprocess technologies involving mixed culture display, high socio-economic interest. From the production of various chemicals, biological compounds to several environmental applications, such bioprocesses are implemented in many ways. However, despite widely use, complexity of the biological mechanisms involved in is still not unresolved. Nowadays, most of the biotechnological processes involving microorganisms are controlled and steer based on the monitoring of physico-chemical parameters. Microbial communities and underlying biological reactions, main actors of the processes, are considered as a "black box". In order to improve the management of such bioprocesses, it is necessary to elucidate these biological "black boxes". Particularly, key questions such as "Who is there?", "Who is doing what?" and "Who is doing what with whom?" should be answered for better understanding of communities' activities. Microbial resource management, a concept previously developed by Verstraete et al. (2007), aims to control and/or steer microbial community capacities through answering these questions.

In this work, microbial resource management has been considered for improving cellulose anaerobic digestion capacities of communities involved in industrial/environmental bioprocesses. Bioaugmentation of two communities with a cellulolytic community isolated from compost has been carried out. Hydrolysis step is indeed recognized as the limiting step in global anaerobic digestion process of cellulosic substrates. For both cases, management of the microbial communities led to an increase of biogas production from (ligno-)cellulosic

substrates. Furthermore, two culture independent techniques have been considered for the assessment of the bioaugmentation treatment applied to leachate communities. Community structure has been monitored through 16s rRNA gene sequencing and flow cytometry. Flow cytometry fits particularly well for a fast and routine monitoring of microbial communities. However, data treatment/transformation is required for an efficient comparison of flow cytometric patterns. Based on this statement, we developed a flow cytometric fingerprinting method allowing the transformation of cytogram into vector of values for further statistical analyses. Both implemented fingerprinting methods give same evidence about microbial dynamics throughout cellulose anaerobic digestion assay. There are however a lack of significant correlation between cytometric and amplicon sequencing fingerprint at the genus or species level. Same phenotypic profiling of microbiota during assays matched to several 16S rRNA gene sequencing ones. Flow cytometry fingerprinting can thus be considered as a promising routine on-site method, suitable for the detection of stability/variation/disturbance of complex microbial communities involved in bioprocesses.

Finally, flow cytometry fingerprinting has been applied for the monitoring of metabolic heterogeneities among monospecies biofilm. Indeed, biofilm is the main form of microbial communities encountered in natural and engineered environment. The development of specific technique aiming at analyzing function of microbial population must then take into account this kind of structure. High degree of specialization can be observed among the cells embedded in biofilm. In order to simplify the approach, the efforts have been focused on single species biofilm in the context of this thesis. Indeed, even isogenic population of cells can take advantage from phenotypic heterogeneities through a division of labor strategy. Such multicellular communities must be considered as a heterogeneous, "multi phenotypes", communities and not anymore as a homogeneous community.

Therefore, similar key questions than for multispecies communities must be answered for elucidating their organization and allowing further efficient management/steering of biofilm based bioprocesses. The efficiency of flow cytometry fingerprint for monitoring metabolic heterogeneities has been proved through a set of experimentations carried out with *Bacillus amyloliquefaciens* as reference organism. Metabolic flow cytometry fingerprinting allows for evaluating the impact of genetic mutation (*B. amyloliquefaciens* mutants were constructed in order to obtain strains with deficient in lipopeptide synthesis) on metabolic profiles.

In conclusion, the management of microbial communities, through bioaugmentation treatment, leads to an improvement of the performances of anaerobic digestion bioprocesses. Moreover, a fast, cheap, culture and operator independent monitoring technique, able to provide crucial information about dynamics of heterogeneous microbial communities has been developed. Efficient for the monitoring of multispecies communities such as the ones involved in anaerobic digestion processes, this approach also reveals its potentialities for the monitoring of phenotypic heterogeneities among isogenic biofilm communities.

Remerciements

Comme tous ceux qui me connaissent le savent, je ne serai jamais un grand écrivain et possède une inspiration littéraire assez limitée. C'est la raison pour laquelle ces remerciements seront assez concis et manqueront cruellement d'originalité.

Je souhaite tout d'abord remercier le Professeur Philippe Thonart de m'avoir proposé de réaliser une thèse au sein de l'unité de bio-industries et madame Destain pour son aide et son soutien moral au cours de la préparation au FRIA ainsi que pour ses conseils lors des premières années de ma thèse.

Je souhaite ensuite remercier mon second promoteur le Professeur Frank Delvigne qui, lorsque je traversais une zone de fortes turbulences, m'a recueilli au sein de son équipe et m'a ensuite guidé jusqu'au bout de cette aventure parsemée d'embuches. En parallèle, je souhaite également te remercier, Frank, pour les bons moments extra-professionnels et notamment houblonnés passés ensemble.

Un tout grand merci à Phidias pour toutes ces agréables promenades sur la décharge de Mont-Saint-Guibert (qu'il pleuve, qu'il vente ou qu'il neige), pour toutes ces discussions footballistiques passionnantes (à quand une virée au Parc des Princes ?) mais aussi pour ton aide précieuse lors des manipulations et rédactions.

Je tiens également à remercier Bernard Taminiau pour son aide dans les travaux de séquençage.

Vu qu'il n'y a pas que le travail dans la vie, je souhaite également remercier tous mes collègues qui ont égayé ces 5 années et plus particulièrement:

- Ceux qui auraient pu arriver plus tôt pour que ces 5 années soient encore meilleures : Thibaut, Auré, Antho

- Sam et John pour vos métaphores plus subtiles les unes que les autres et surtout « vos encouragements » (dixit le plus mauvais de la fac).

- Olivia pour toutes ces discussions enrichissantes et passionnantes.

- Quentin pour toutes les discussions constructives (ou pas) autour d'une (ou plusieurs) bonne(s) bière(s)

- Thanh pour ta positive attitude et ton éternelle bonne humeur. On se boit bien vite une Chouffe et une Chimay bleu chez toi au Vietnam.

Enfin un énorme merci à ma famille et amis et tout particulièrement à toi Marie, qui m'a toujours soutenu et remotivé lors des (nombreux) moments de doute que j'ai pu traverser.

List of publications

Article 1 (included in Chapter I)

Kinet, R., Destain, J., Hiligsmann, S., Thonart, P., Delhalle, L., Taminiau, B., Daube, G., Delvigne, F., 2015. Thermophilic and cellulolytic consortium isolated from composting plants improves anaerobic digestion of cellulosic biomass: Toward a microbial resource management approach. Bioresour. Technol. 189, 138–144. doi:10.1016/j.biortech.2015.04.010

Article 2 (included in Chapter II)

Kinet, R., Dzaomuho, P., Baert, J., Taminiau, B., Daube, G., Nezer, C., Brostaux, Y., Nguyen, F., Dumont, G., Thonart, P., Delvigne, F., 2016. Flow cytometry community fingerprinting and amplicon sequencing for the assessment of landfill leachate cellulolytic bioaugmentation. Bioresour. Technol. doi:10.1016/j.biortech.2016.04.131

Article 3 (included in Chapter III)

Kinet, R., Ongena, M., Boon, N., Delvigne, F., 2016. Dynamic of biofilm formation by *Bacillus amyloliquefaciens* deciphered by flow cytometry phenotypic fingerprinting. Will be submitted in Microbial Bitoechnology

Article 4

Baert, J., **Kinet, R.**, Brognaux, A., Delepierre, A., Telek, S., Sørensen, S.J., Riber, L., Fickers, P., Delvigne, F., 2015. Phenotypic variability in bioprocessing conditions can be tracked on the basis of on-line flow cytometry and fits to a scaling law. Biotechnol. J. 10, 1316–1325. doi:10.1002/biot.201400537

List of abbreviations

MCB	Mixed culture biotechnology
MRM	Microbial resource management
PCR	Polymerase in chain reaction
ARDRA	Amplified ribosomal DNA restriction analysis
t-RFLP	Terminal-restriction fragment length polymorphism
RISA	Ribosomal intergenic spacer analysis
RAPD	Random amplification of polymorphic DNA
DGGE	Denaturing gradient gel electrophoresis
TGGE	Temperature gradient gel electrophoresis
OTU	Operational taxonomic unit
RDP	Ribosomal database project
LH-PCR	Length heterogeneity polymerase in chain reaction
SSCP	Single strand conformation polymorphism
Rr	Range weighted richness
Co	Community organization
Fo	Functional organization
Ep	Ecological pareto
Cd	Community distorsion
Dy	Dynamics
FCM	Flow cytometry
GFP	Green fluorescent protein
FISH	Fluorescence in situ hybridation
FSC	Forward scatter
SSC	Sideward scatter
RSG	Redox sensor green
DAPI	4',6'-diamidino-2-phénylindole
CHIC	Cytometric histogram image comparison
CyBar	Cytometric Barcoding
2D	2 dimensions

PB	Probabilitybinning
2D-GE	2 dimensional gel electrophoresis
DIGE	Difference in gel electrophoresis
MS	Mass spectrometry
LC-MS	Liquid chromatography – mass spectrometry
GC-MS	Gas liquid – mass spectrometry
CE-MS	Capillary electrophoresis – mass spectrometry
VFA	Volatile fatty acids
AO	Acridine orange
DNA	Deoxy ribonucleic acid
RNA	Ribonucleic acid
mRNA	Messenger ribonucleic acid
ACE	Abundance-based coverage estimator
MSW	Municipal solid waste
CETeM	Centre d'enfouissement technique de Mont-St-Guibert
COD	Chemical oxygen demand
WWTP	Waste water treatment plant
NCBI	National center for biotechnology information
DMSO	Dimethyl sulfoxide
PBS	Phosphate buffered solution
UPLC	Ultra performance liquid chromatography
HPLC	High performance liquid chromatography
OD	Optical density
EPS	Exopolysacchararides

Table of content

1 Introduction

1.1 Monitoring of microbial population involved in bioprocesses

Microbial biotechnologies have a strong socio-economical potential. From the production of various fine chemicals, enzymes and therapeutically compounds to several environmental "application", bioprocesses can be implemented in many ways. However, despite widely use, complexity of the biological mechanisms involved in these processes is still not unraveled. The microorganisms and underlying biological reactions, main actors of the processes, are considered as a "black box". Most of the biotechnological processes involving microorganisms are controlled and steer based on the monitoring of bulk physico-chemical parameters. In order to achieve optimal exploitation of the particular microorganisms, it is necessary to elucidate these biological "black boxes". Particularly, the structure of microbial communities should be deeply characterized. Improving monitoring capabilities is thus crucial to optimize bioprocesses.

Traditionally, culture dependent methods, such as heterotrophic plate count, were used for controlling performance of biotechnological processes or for investigating microbial population diversity. However, as well for natural samples and axenic cultures monitoring, these techniques exhibit several drawbacks:

- Incubation periods are unavoidable, leading to a significant lag time between the analysis and the delivery of the results.

- Majority of bacterial species are uncultivable. 99% of microorganisms observable in nature typically are not cultivated using standard techniques (Amann et al., 1995).

- Cellular viability assessment is uncertain. Species in a viable but non cultivable state are ignored.

- Single-cell analysis is not possible, observation are done at colonies level. Genotypic and/or phenotypic variability affecting cells forming a colony are not considered

- Metabolic network studies are restricted to the observation of "On/Off" responses.

In order to cope with these restrictions, culture independent methods, such as "-Omics techniques" and flow cytometry, have been largely developed during the last years. Omics techniques provide deep understanding of biological mechanisms through the monitoring of the diverse molecular intermediates involved in (proteins, DNA, RNA, …) (Zhang et al., 2010). On the other hand, flow cytometry allows highlighting microbial heterogeneity at single cell level in a very short time through the measurement of a set of optical parameters (Delvigne et al., 2014; Delvigne and Goffin, 2014; Koch et al., 2013). Flow cytometry allows on-site routine microbial communities analyses required for efficient pro-active management of bioprocess while "-Omics" allows detailed analysis. Based on the efficiency of culture independent methods for elucidating communities' complexity, Verstraete et al. (2007) developed the concept of Microbial Resource Management (MRM). MRM aims to solve practical problem through the use of microorganisms (Read et al., 2011). More particularly, solutions are supplied through control and/or steer of microbial community capacities. Especially, bioaugmentation technology, in a microbial resource management view, consists in adding an exogenous microbial community with specific activity in order to improve the performance of the in-situ biological processes, and facilitate the establishment of specific species in microbial communities (Bouchez et al., 2000).

In the environment, many microbial communities exhibit biofilm organization. The development of monitoring technique aiming at improving the management of microbial populations must then take into account this kind of structure. Actually biofilm communities are largely investigated with advanced imaging approach (e.g. confocal laser scanning microscopy) and microfluidic devices (Bruchmann et al., 2015; Morgenroth and Milferstedt, 2009). However, these techniques only evaluate the biofilm structure at microscale. Efforts still must be done for linking these microscales investigations with meso/macroscale observations. This link is crucial as mechanisms and processes at the microscale depend on and influence interactions on larger spatial scales (Morgenroth and Milferstedt, 2009).

1.2   Objectives of the thesis

The objective of the thesis is the elucidating of microbial complexity associated to the anaerobic digestion of cellulose in order to improve process efficiency. Elucidating microbial communities' complexity is crucial for their efficient management. "-Omics" techniques are suitable for this purpose. However, these techniques are time consuming and required expensive and sophisticated materials. On the opposite, flow cytometry displays prerequisites for a routine communities' structure analysis: cheap, fast and simple implementation. Based on this statement, we will tackle on the development of a flow cytometry based operator-independent method allowing high densities communities monitoring. Moreover, as many microbial communities involved in environmental biotechnologies form biofilm structure, efforts will be done for developing a similar flow cytometry monitoring approach taking into account the particularity of such multicellular communities.

1.3 Refernces

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

Bouchez, T., Patureau, D., Dabert, P., Wagner, M., Delgenès, J.P., Moletta, R., 2000. Successful and unsuccessful bioaugmentation experiments monitored by fluorescent in situ hybridization. Water Sci. Technol. 41, 61–68.

Bruchmann, J., Sachsenheimer, K., Rapp, B.E., Schwartz, T., 2015. Multi-channel microfluidic biosensor platform applied for online monitoring and screening of biofilm formation and activity. PloS One 10, e0117300. doi:10.1371/journal.pone.0117300

Delvigne, F., Goffin, P., 2014. Microbial heterogeneity affects bioprocess robustness: Dynamic single-cell analysis contributes to understanding of microbial populations. Biotechnol. J. 9, 61–72. doi:10.1002/biot.201300119

Delvigne, F., Zune, Q., Lara, A.R., Al-Soud, W., Sørensen, S.J., 2014. Metabolic variability in bioprocessing: implications of microbial phenotypic heterogeneity. Trends Biotechnol. 32, 608–616. doi:10.1016/j.tibtech.2014.10.002

Koch, C., Gunther, S., Desta, A.F., Hubschmann, T., Muller, S., 2013. Cytometric fingerprinting for analyzing microbial intracommunity structure variation and identifying subcommunity function. Nat Protoc 8, 190–202. doi:10.1038/nprot.2012.149

Morgenroth, E., Milferstedt, K., 2009. Biofilm engineering: Linking biofilm development at different length and time scales. Rev. Environ. Sci. Biotechnol. 8, 203–208. doi:10.1007/s11157-009-9163-1

Read, S., Marzorati, M., Guimarães, B.M., Boon, N., 2011. Microbial Resource Management revisited: successful parameters and new concepts. Appl. Microbiol. Biotechnol. 90, 861–871. doi:10.1007/s00253-011-3223-5

Verstraete, W., Wittebolle, L., Heylen, K., Vanparys, B., de Vos, P., van de Wiele, T., Boon, N., 2007. Microbial Resource Management: The Road To Go for Environmental Biotechnology. Eng. Life Sci. 7, 117–126. doi:10.1002/elsc.200620176

Zhang, W., Li, F., Nie, L., 2010. Integrating multiple "omics" analysis for microbial biology: application and methodologies. Microbiology 156, 287–301. doi:10.1099/mic.0.034793-0

2 State of the art

2.1 Mixed culture biotechnology

Nowadays, most of biotechnological industrial processes are based on pure culture of isolated microorganisms. However such strategy is contradictory with strategy involved in the majority of biogeochemical processes led by microorganisms on Earth. Microorganisms generally exist in association (e.g. with other microbes or other type of organisms) and implement cooperative strategies (cross feeding, aggregation and adhesion).

Mixed culture biotechnology (MCB) exploits such highly organized natural microbial communities for bioprocesses fitting in very wide field of applications (e.g. energy production, wastewater treatment, sustainable recovery of resource, biobased chemicals …). Positive interaction between members of consortia allows accomplishing impossible tasks for isolated strains. Moreover, in some cases, MCB could overcome traditional pure culture based bioprocesses thanks to several specific advantages such as (1) no sterilization requirements, (2) adaptive capacity owing to microbial diversity, (3) the capacity to use mixed substrates, and (4) the possibility of a continuous process (Kleerebezem and van Loosdrecht, 2007).

Anaerobic digestion is a perfect example of the usefulness of natural microbial communities for facing different industrial and environmental challenges. Coordinated action of manifold different species, displaying specific metabolic activities, allows valorization of complex substrates (usually considered as waste products) as a source of energy as well as a source for building block molecules production.

However, efficient management of processes involving consortia is not easy which restrains an expanded use of MCB. The complexity of metabolic network and relationships between members induces difficulty in controlling/steering of the community. Actually, the majority of processes involving microbiome are operated according to the measurement of bulk chemical, physical and biological parameters (temperature, enzymatic activity, substrate and products concentrations, …) (Koch et al., 2014b). A posteriori regulation is implemented during processes. Conditions adjustment follows detection of bulk parameters disturbance. These parameters do not allow proactive management as no information are available for segregated analysis of the respective actors' performances which is necessary for predicting performance of microbiome reactor.

Hence development of a methodology for direct monitoring of bioprocesses actors and for elucidating their respective participation to overall performance is necessary.

2.1.1 Elucidating microbiome blackbox – Technical requirements

Based on the statement that domesticated natural communities can provide great services in different context, Verstraete et al. (2007) developed the concept of microbial resource management. Microbial resource management (MRM) concept aims to solve practical problem through use of microorganisms (Read et al., 2011). More particularly, solutions are supplied through the control and/or steering of microbial community capacities. However, in order to provide these solutions, MRM must include tools for elucidating microbiome blackbox.

In this context, three key questions must be answered:

(1) Who is there?

(2) Who is doing what?

(3) Who is doing what with whom?

2.1.1.1 Who is there?

2.1.1.1.1 Molecular approaches

Molecular fingerprinting techniques

Widely used in microbial ecology, molecular fingerprinting techniques are perfect to unravel complexity of communities and to answer above mentioned key questions (Verstraete et al., 2007). They provide a pattern or profile of the genetic diversity in a microbial community.

DNA extracted from samples is involved as it is for these techniques. However, only a part of the total genetic information is analyzed. Classically, genome sequences of interest correspond to the genes of the ribosomal operon, and particularly the *rrs* gene (16S rDNA) (Ranjard et al., 2000). Genetic fingerprint is obtained after Polymerase in Chain Reaction (PCR) which is applied for amplification of the targeted regions. According to the technique applied, diversity of amplified sequences is resolved by differential electrophoretic migration on agarose or polyacrylamide gels, based on their size (ARDRA, t-RFLP, RISA, RAPD) or their sequence (DGGE,TGGE) (Ranjard et al., 2000).

However, such techniques display some drawbacks. One of the most preoccupant is their lack of sensitivity (e.g. detection limit of DGGE technique is 1% of the total DNA (Marzorati et al., 2008)). Another limitation is relative to the resolution of the bands constituting the patterns. Normally, a particular species can be associated to each band of the pattern as it corresponds to a particular sequence. However, particularly with ribosomal sequences, one band can be associated to several species and inversely. So interpretation of the diversity can be biased.

Next generation sequencing techniques

Recent development of high-throughput sequencing techniques provides new tools to overcome limitations associated to molecular fingerprinting techniques and to investigate more deeply the structure and the functionality of complex communities. Most of these techniques rely on sequencing-by-synthesis design and require preliminary DNA amplification step (PCR) (Di Bella et al., 2013). Previously, first generation sequencing relied on Sanger sequencing method (chain termination method) (Sanger et al., 1977).

When applied in order to elucidate the composition of a microbial community, sequencing techniques target only a part of genetic information extracted from the sample, usually corresponding to one or more fragment of 16S rRNA gene. Amplicon sequencing is the term used to express such approach. Choice of the variable region(s) of 16S rRNA gene to analyze is not easy as inter-species variability appears in different ways for these regions. Choice among nine variable regions must take into account information specific to samples. Notably, species diversity targeted and level of resolution affect this choice (Di Bella et al., 2013). Despite efficiency of 16s rDNA sequencing for elucidating structure of a microbiome, it is important to be aware of several limitations inherent to the use of this ribosomal gene. Due to a high level of sequence divergence between multiple copies of the gene, different 16S rDNA sequences from a given bacterium can result in two or more distinct Operational Taxonomic Units (OTUs) what induces overestimation of microbial diversity in a sample. Moreover, multiple copies of the rrn operon, which contains the 16S rRNA gene, displayed by some bacteria induce over-representation of these bacteria when investigating abundances. Lastly, some 16S rDNA primers, although considered universal, preferentially bind to some taxa over others, thus over-representing those taxa in experiments (Di Bella et al., 2013). Hence, if next generation sequencing also provides quantitative information

through relative abundance of different OTUs, it must be considered with caution as different steps of the protocol can induce bias in analysis.

Logically first step of high-throughput methods is the microbial genomic DNA extraction-purification. As microbes can display wide range of structure and composition according to their environments, DNA extraction protocol must be judiciously selected (Di Bella et al., 2013). This step is potentially a source of bias for quantitative analyses as preferential DNA extraction occurred according to the microbe's structure.

Following extraction, DNA samples are prepared for ulterior steps. Despite dissimilarity according to the type of sequencing techniques, this preparation includes three similar steps: (1) fragmentation of DNA molecules for providing sequences that can be further analyzed, (2) addition of blunt ends to aid further processing and (3) ligation of adaptors (platform-specific) to the fragments for acting as amplification primers and for allowing the fragments to be attached to solid surface for sequencing (Di Bella et al., 2013).

When attached to solid surface, DNA sequences are amplified to form clone libraries. The solid surface, where clone libraries are formed, differs according to the system employed. Some involve microbeads as PCR support (Roche 454, Ion Torrent PGM) whereas flow cell are implemented in other system (Illumina platform). These supports lead respectively to microbeads covered of identical DNA fragment after emulsion PCR and to clusters of identical DNA fragments after "PCR bridge". Once clone libraries are formed, they can be sequenced. Again, sequencing protocol vary according to respective systems. Some rely on optical signals for evidencing the incorporation of respective nucleic bases (454 Roche and Illumina) while Ion Torrent platform is based on direct detection of protons released linked to base incorporations.

High throughput sequencing produces very large amount of data. Computational analysis with different informatics tools is necessary to get any useful information from these data. In term of targeted amplicon sequencing, comparison of acquired sequences with existing databases results in identifying who is there and in which proportion. However, preliminary treatment steps must be applied to raw sequencing data. Typically, required bioinformatics tools are grouped together into packages. Among these ones, MOTHUR (Schloss et al., 2009) is one of the most popular. This package aligns and clusters the sequences while algorithms such as Pyronoise and UCHIME (Edgar et al., 2011) are used for denoising and chimera detection. After data treatment, comparison with existing database such as SILVA or Ribosomal Database Project (RDP) provides taxonomical assignation.

MRM parameters extracted from genetics analyses

Interpretation of data resulting from these different methods is quiet ambiguous. Based on this statement, practical tools (MRM parameters) have been designed (Marzorati et al., 2008; Read et al., 2011). They allow for quantitative and universal description of community structure. Originally, tools have been designed to fit with DGGE patterns processing but rapidly their use has been extrapolated to other classical fingerprinting molecular techniques (LH-PCR, tRFLP, TGGE, SSCP, and Clone Libraries). Furthermore, recent development and widespread use of high throughput methods require the adaptation of originally designed parameters and even the design of new ones.

Read et al. (2011) give non exhaustive list of studies in which MRM tool set successfully describe structure of community in different environments and through use of these molecular fingerprinting methods.

Range-weighted richness (Rr)

The range-weighted richness (Rr) reflects the microbial diversity of a communities and is originally calculated according to the following formula (1) for DGGE analysis:

$$Rr = (N^2 \times D_g) \qquad (1)$$

Where N represents the total number of bands in DGGE pattern and D_g the denaturing gradient comprised between the first and the last band of the pattern (Marzorati et al., 2008). The denaturing gradient value is the percentage needed to represent the total variability of the genetic content present in the sample. This value is directly linked to the nature of the analyzed environment. More an environment is able to host microorganisms (i.e. more environment display favorable conditions), more required gradient is wide (i.e. more genetic diversity is wide).

The observations of values calculated for different environments (referenced by Marzorati et al. (2008)) put Rr values in perspective. The environments displaying unfavorable conditions (e.g. polluted soil (Marzorati et al., 2005), deep sea hydrothermal site (Postec et al., 2005)) are characterized by Rr value <10; they present low diversity. Intermediate values, i.e. between 10 and 30, can be considered as medium range-weighted richness while values superior 30 correspond to high range-weighted richness, characteristic of environment such as garden soil (Edenborn and Sexstone, 2007) able to host several species with wide range of genetic content. In these cases, Rr parameter established with classical molecular fingerprinting methods quantifies richness and genetic diversity of diverse communities. However, values are dependent towards technique applied and its detection threshold. Bent et al. (2007) highlighted while classical fingerprinting methods are useful for comparative analyses, they strongly underestimate richness of complex communities.

New 16S rDNA based sequencing methods successfully highlight previously unseen bacterial species as they are more sensitive and they cover more completely taxonomic diversity. In respect to this, Read et al. (2011) redefined the way to calculate *Rr* parameter when sequencing method is carried out. The value should include both the number of unique OTUs, estimated by rarefaction curves analyses, and phylogenetic broadness of the present OTUs on family, class or order level. Rarefaction curves are constructed by plotting the number of OTUs (y-axis) against the corresponding number of sequence analyzed (to obtain this number of OTU). The asymptote of the curve provides estimated maximal number of OTU present in the sample.

Community organization (Co)

After elucidating the extent of the biodiversity reservoir, it is necessary to highlight the respective participation of each member to community functionality. In its definition of MRM, Verstraete et al. (2007) maintain that Pareto principle, which in respect to market economy tables that 80% of the "goods" are in the hands of 20% of the concerned population, is applicable to microbial ecology. Among overall species distribution, 20% of the species control 80% of the energy flux (Verstraete et al., 2007). With this in mind, functional organization (Fo) parameter was designed in order to characterize the ability of a community to organize in a such way that functionality is ensured. Structure and functionality of the community are correlated through this parameter. However, it rapidly appears that a link between a given functionality and a respective group of microorganisms cannot always be established. Based on this statement, Fo parameter has been renamed as Community organization (Co), a parameter which only describes the microbial community in terms of evenness degrees (Read et al., 2011).

Species distribution inside a community at a given time (i.e. in given environmental conditions), is graphically represent through the construction of Pareto-Lorenz evenness curves. These curves are constructed by plotting cumulative proportion of OTUs (x-axis) against cumulative proportion of OTU abundances (y-axis) (Figure 1).
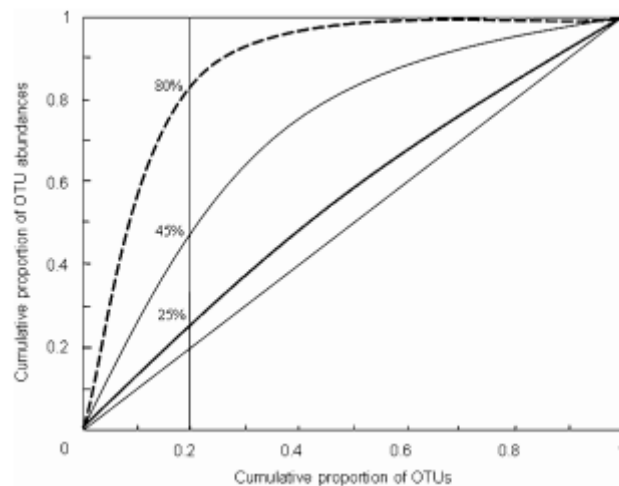
Figure 1 Pareto-Lorenz curves derived from three hypothetical DGGE patterns. The 25%, the 45% and the 80% curves refer to a low, medium and high functional organization respectively. The 45° diagonal represents the perfect evenness of a community. (Marzorati et al., 2008)

On such representation, 45° diagonal represents a perfect evenness, i.e. each member of the community displays the same relative abundance. Contrary, more a curve differs from this reference line, less this equitable distribution is respected. In such case of weak evenness, communities are dominated by a small fraction of the detected species while the rest of the population, displaying weak relative abundance, is waiting for more favorable conditions. Figure 1 displays three theoretical community organizations. The percent values associated to each curve do not represent Co value but the score from the y-axis projection of curve interception with vertical 20% x-axis line. Attributing these scores to curves was suggested by Wittebolle et al. (2008) for more easy interpretation and comparison. More this value is high, more community is uneven. 25%, 45% and 80% curves represent community with respectively low, medium and high unevenness.

13

Co parameter is calculated as the Gini coefficient times 100 with Gini coefficient expressing the normalized area between distribution curve and perfect evenness diagonal. Co has a percent value varying from 0 to 100. Furthermore, as Gini coefficient is inversely proportional to degree of evenness, high Co values correspond to high unevenness (i.e. 0=even community, 100 =uneven community).

As previously mentioned, amplicon sequencing highlights all species from a community, even those displaying very weak relative abundance. This previous unseen majority alters the shape of Lorenz curve and thus Gini coefficient. Consequently, new way to graphically express the evenness of a community is necessary when sequencing methods are used (Read et al. (2011)). Consider a dataset of *N* OTU classes:

1° OTU classes are sorted from high to low abundances (with OTU-1 the most abundant and OTU-N the less)

2° Co values (Co^i, with *i*=1 to *N*) are calculated for a progressing window of OTUs from 1 to N

3° Co^i are plotted against the amount of classes (i) to obtain the descriptive curves.

Thanks to these curves, it is possible to determine when microbial community is reaching its maximum Co value which represents its degree of evenness.

In relation with Pareto law, it is possible for determining the optimal Co value for a given ecosystem. This value is called Ecological Pareto (Ep) and corresponds to the optimal community organization in a given environment (Read et al., 2011). Furthermore, Community distortion (Cd) parameter has been created for assessing deviation of a community from the optimal structure. Absolute values describe the severity of distortion while +/- indicates, respectively, a too uneven or a too even population (Read et al., 2011).

Dynamics (Dy)

To maintain functionality, communities must adapt in response to environmental modifications. Consortia must be considered as a continuum of cooperative communities succeeding each other in case of unregulated environments. Based on this statement, dynamics (Dy) parameter has been created (Marzorati et al., 2008). It describes to what extent a community is able to change its structure in a given time period. Practically, it represents the number of species which can be detected during a given laps of time. More species can be detected; more the community is dynamics and more the original community may evolve (Marzorati et al., 2008). Consequently, this parameter is highly technique dependent. Comparison of Dy values resulting from studies involving different methods is not possible. To overcome this drawback, it is preferable to deal with the rate of change affecting a community during a given period. The rate of change averages the degree of change between successive fingerprints of a community over a fixed time interval (Marzorati et al., 2008). In the case of DGGE fingerprinting, % of change between samples are calculated according to equation (2):

$$\% \, change = 100 - \% \, similarity \quad (2)$$

Where % of similarity result from Pearson product-moment correlation coefficients calculated on basis of densiometric curves of different patterns.

Logically, a high rate of change corresponds to a community displaying a high level of dynamics while a low rate is linked to a nearly stable community. For example, Wittebolle et al. (2005, 2008) highlight different levels of dynamics in several ammonia-oxidizing bacterial communities from different reactors (based on DGGE patterns). First community, characterized by a rate of change of 3% (Wittebolle et al., 2005), can be considered as

weakly dynamic. Community from the second reactor, displaying a rate of change of about 12% is assumed to be moderately dynamics. Finally, a highly dynamic population is observed in the last reactor in which rate of change of approximatively 25% is calculated.

Contrary to previous parameters, Dy does not need adaptation to be compatible with sequencing methods. As for DGGE patterns, differences between samples are easily evaluated thanks to Pearson product-moment correlation coefficients.

Alternative parameters

In parallel to MRM parameters, alternative parameters can be calculated based on sequencing data for evaluating the richness and the diversity of a community (Di Bella et al., 2013). Richness and more precisely the minimum number of OTU present in a community can be estimated by Chao1 (Chao, 1984) and ACE (Chao and Lee, 1992) parameters. Furthermore, community diversity which combines richness and evenness is traditionally expressed through alpha-diversity indices and more precisely the Shannon or Simpson diversity indexes (see Li et al. (2012) for more a detailed description).

2.1.1.1.2 Flow cytometry for routine analysis of microbial communities in bioprocesses

Above described molecular techniques are powerful methods for elucidating natural communities in term of composition. Nonetheless, several drawbacks are still inherent to these ones. In term of bioprocesses monitoring, major disadvantage is the incompatibility of such techniques with on-site routine analysis as they require expensive specialized equipment and are relatively high time-consuming for data acquisition and interpretation. As some bacterial species are characterized by generation times of minute and so can rapidly evolve in response to environmental modifications, a dense population sampling and analysis is required for an efficient community monitoring. Flow cytometry combined with

biostatistics based tools (flow cytometry fingerprinting) is a promising approach to achieve this goal and so enabling efficient process control/steering.

Flow cytometry

Flow cytometry (FCM) is a high throughput technique for single-cell analysis. The analysis is based on measurement of cellular optical characteristics.
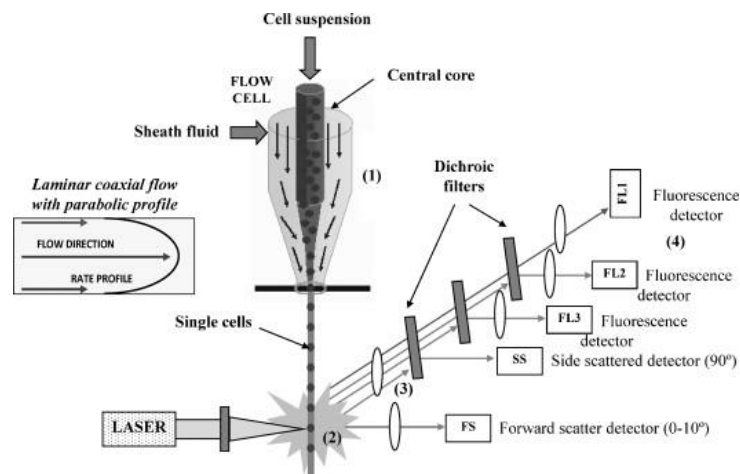
Figure 2 Flow cytometry measurement principle (Díaz et al., 2010)

FCM measuring principle can be divided in five steps:

(1) Cells are organized into a single stream of particles thanks to hydrodynamic focusing.

(2) Single cells pass successively through laser beam what induces emission of signals related to cells' optical characteristics. When crossing cells, laser induces two types of interaction: (1) light scattering resulting from cellular intrinsic characteristics (size, internal organization, and granularity) and (2) emission of fluorescence which can result either from cell auto-fluorescence either from fluorochromes (fluorescent dyes or fluorescent tag such as GFP or FISH approach).

(3) Scattered and fluorescent signals are separated thanks to a set of filters and mirrors.

(4) Detection of different signals by photodiodes. Two are attributed for the measurement of scattered signals; low angle (below 10°) deviation from incident light is called Forward Scatter (FSC) while Sideward Scatter (SSC) is associated to perpendicular deviation. The numbers of fluorescent detectors vary according the to complexity of material and the related working wavelength spectrum (FL-1, FL-2, FL-3 on figure above).

(5) Optical signals are transformed into electronical signals enable to be digitalized and finally analyzed by software.

Through the measurement of a set of optical parameters, FCM is so able to characterize cell population at single-cell level in a very short time. The characterization is based as well on intrinsic as on extrinsic optical parameters that can be related to structural and/or functional cell properties. FSC and SSC typically give structural information about respectively cell size (light deviation increases with cell size increase) and about internal cell organization. Nevertheless, these information are generally not enough powerful for efficient subpopulation discrimination among a global population. Contrary, through highlighting supplemental physiological or functional cell's properties, fluorescent labelling allow for achieving this goal. According to cell target sites, different probes are used. These probes can be of different nature: (1) fluorescent dyes which increase fluorescence after specific binding to cellular molecules or components, (2) fluorescent dyes which accumulate in specific compartment of cells, (3) fluorescent dyes whose fluorescence is modulated through environmental conditions or enzymatic activity, (4) Green Fluorescent Protein (GFP) genetically-encoded in cells or (5) fluorochromes conjugated to antibodies or oligonucleotide (FISH). Membrane integrity is for example easily estimated thanks to propidium iodide (PI) staining. Dye penetrates and links to DNA of cells displaying damaged membrane while it

cannot penetrate cell with intact membrane (Díaz et al., 2010). For their part, Redox Sensor Green (RSG) dye highlights actively metabolizing microorganisms (Kalyuzhnaya et al., 2008) and tetracycline hydrochloride highlights poly-phosphate accumulation in cell (Günther et al., 2012) . Cellular DNA content is also a suitable target for community analysis and cell cycle study. Sybr Green I and Syto9, who stain indifferently entire nucleic acid content, and 4',6'-diamidino-2-phénylindole (DAPI), who links specifically to adenine and thymine nucleic bases, segregate subcommunities according to their respective DNA content (or A/T content). For more examples and a detailed reviewed of fluorescent staining techniques, see Díaz et al. (2010).

Flow cytometry fingerprinting of microbial communities

The resulting sets of optical parameters acquired for each cell during FCM measurement are usually represented through biparametric cytometric histograms (cytogram) (figure 3). Such representation induces the formation of different clusters containing cells displaying similar characteristics for the chosen parameters. Similarly to DGGE patterns, these histograms can be considered as fingerprints of a microbial population at a given moment after mathematical data treatment (figure 3B). Fingerprinting approach facilitate the interpretation of cytometric histograms and more particularly for parametrizing differences between samples (Kinet et al., 2016). Informatics tools have been developed for constructing these fingerprints. Koch et al. (2014) compare the different tools available for interpretation of communities cytogram: Dalmatian Plot, Cytometric Histogram Image Comparison (CHIC), Cytometric Barcoding (CyBar) and FlowFP. Each of these tools displays specific way to work.
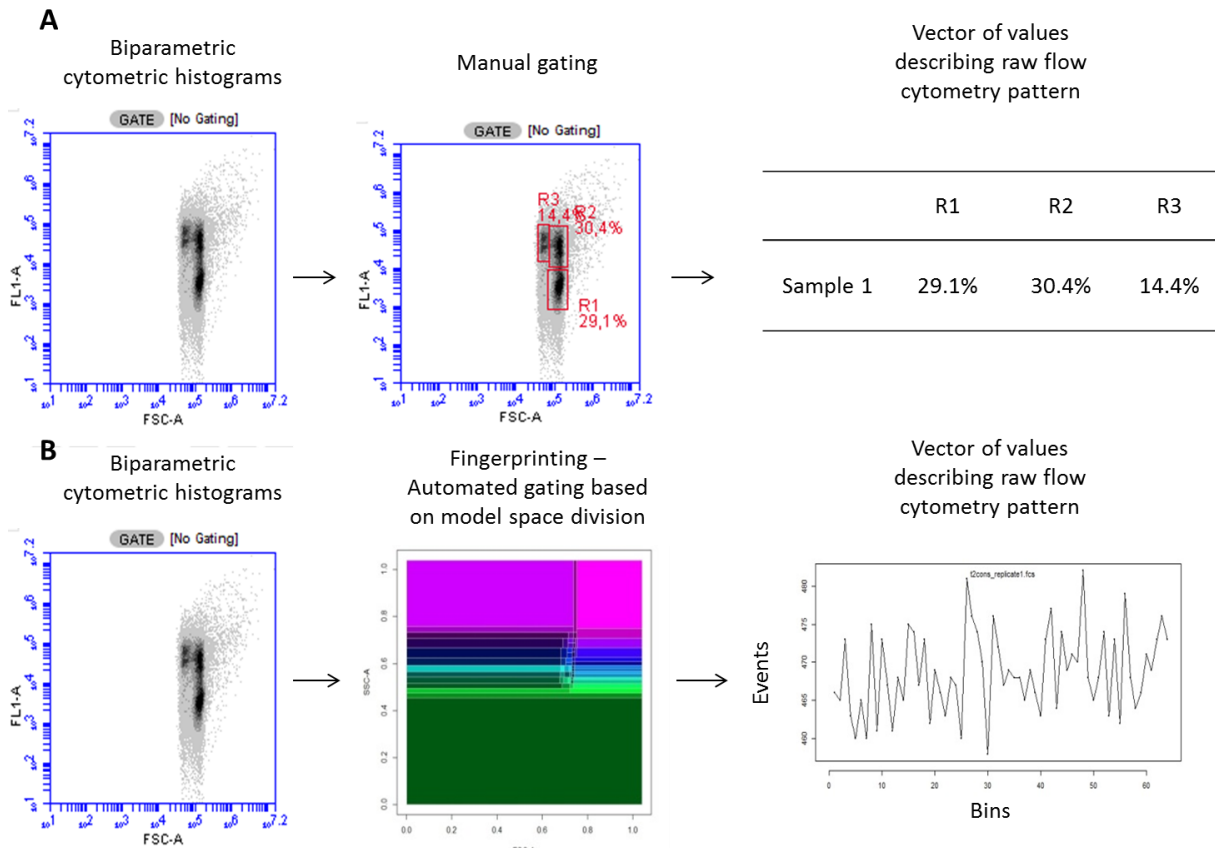
Figure 3 Analysis of raw biparametric cytometric histograms (FSC-A/FL1-A). A vector of values describing raw flow cytometry pattern is obtained through either (A) a manual gating procedure or (B) an automated gating procedure (FlowFP fingerprinting). The automated gating procedure depicted here involved the establishment of a grid composed of 64 bins. The structure of this grid depends on the structure of the raw biparametric cytometric histograms studied. Contrary to procedures involving manual gating, this fingerprinting technique is operator-independent.

Dalmatian Plot

Bombach et al. (2011) were the firsts to describe this technique involving simplified black-and-white images for representing communities' cytogram. First step of the method consists in gating (with circle) the most abundant subsets of cells for each compared 2D histogram. For each cytogram, this results in black circles, representing the different observed cell clusters, over white back ground. Consequently, only presence/absence information result from this black/white gating. Interestingly, cell abundancies can be evaluated through insertion of a grey scale in circled gates (Bombach et al., 2011; Müller et al., 2012). Then, an image analysis is performed thanks to adequate software and related macros. The gates' areas are determined through the count of contained black pixels. Finally, images from two

samples are overlapped and area corresponding to overlapping gates is quantified. Dissimilarity values result from each paired comparison and are the basis for the construction of a dissimilarity matrix for further statistical analysis.

Cytometric Histogram Image Comparison (CHIC)

Similarly to Dalmatian Plot, CHIC involves image analysis (Koch et al., 2013a). However, it does not require operator gating. Cytometric histograms are directly converted to gray scaled images thanks to adequate cytometric software. Image resolution used to represent histogram varies according to desired precision. Logically, resolution must be similar when comparing samples. Then, images paired comparison is realized thanks to software (construction of overlap and XOR images to measure dissimilarity between images) and matrix dissimilarity is constructed.

Cytometric Barcoding (CyBar)

First step corresponds to a gating step, every cell clusters are marked in each cytogram thanks to ellipsoid gates. Secondly, a gate template is created by combination of all previously designed gates. Finally, this gate template is applied to all samples and number of cells in each gate is calculated for each sample. A matrix with the respective cell abundancies per gate for each sample is obtained (Koch et al., 2013).

FlowFP

FlowFP is a software package developed by Rogers and Holyst (2009) and available at open source Bioconductor platform (Gentleman et al., 2004). It consists in creating a n-dimensional quantitative fingerprint of each sample from bivariate flow cytometry (FCM) distributions using the recursive probability binning (PB) algorithm implemented in the Bioconductor package FlowFP (figure 3B) (Rogers and Holyst, 2009). In a first time, a model,

21

composed of hyper-rectangular regions (bins) of varying sizes and shapes, is established. From superposed FCM distribution (i.e. data from all samples are pooled together) and thanks to PB algorithm (Rogers et al., 2008), bivariate data space is divided in hyper-rectangular regions in such way that each contains similar number of events (one event corresponding to one cell). The first step of the algorithm consists in the division of space into two bins containing similar number of events. Afterwards, each of these bins is again divided into two bins with equal number of events, and so forth. Therefore the regions of bivariate FCM histogram displaying high density of events are characterized by bins of small area whereas larger bins characterized regions of weak density. Moreover, the final number of bins (n) is arbitrary set and correspond to 2^i with i the number of recursive subdivisions. The obtained model is then applied to each sample and number of cells per bin is determined, creating a feature vector of counts (n-dimensional) for each sample. The latter is also referred to as the fingerprint.

Comparison of the fingerprinting methodologies

As underlined above, the different analytical protocols results in different outcomes. Dalmatian plot and CHIC provide dissimilarity matrix based on paired samples comparison while CyBar and FlowFP provide vectors with number of cells per gate/bin for each sample. The two first methods provide trends about population dynamics at whole community level whereas others provide a segregated approach according to gates or bins. Such individual change detection is advantageous as simple statistical tools like principal component analysis can transform segregated vision into whole community approach while the inverse is not possible.

Operator dependency is another important feature. Indeed, operator depending steps, such as gating in Dalmatian Plot and CyBar, induce subjectivity in interpretation and so do not allow comparison of results treated by different operators. Contrary, CHIC and FlowFP are not operator dependent. But these techniques require the adjustment of some options, respectively the cytogram imaging resolution and the number of recursion for template construction. Differently, drawback of FlowFP is the absence of biological meaning in binning procedure contrary to Dalmatian Plot and CyBar whose subcommunities are considered for gating. Consequently, FlowFP is less compatible with cell sorting than CyBar.

Each tool displays advantage and disadvantage. However, according to above mentioned advantages of FlowFP and its ease of use (open source R package), this approach has been considered in the context of this work for the analysis of flow cytometry data.

2.1.1.2 Who is doing what? Who is doing what with whom?

Once the structure of a community is established, it is useful to understand the way it works. Different molecular techniques can provide required information to reach this goal.

First we can cite methods based on the analysis of communities' metagenome. Metagenome term was proposed first time by Handelsman et al. (1998) to describe "genome of total microbial resources found in nature". In other terms, the metagenome represents total genomic information from all microorganisms in a given environment. Differently from amplicon sequencing, metagenome sequencing investigates all genomic information from the population so that all genes and associated potential functionalities are highlighted.

Metatransciptomics, metaproteomics and metabolomics are other useful molecular methods for elucidating complex mechanisms involved in a community. Contrary to genome based approach which provides non dynamic information about potential functionalities,

these techniques allow dynamic analysis of functionalities involved by a community at a given moment. Thanks to these methods, the link between genomic information and functionalities of a community should be established. Indeed all these methods monitor molecular species directly involved in (metatranscriptome and metaproteome) or resulting from (metabolome) the different cellular mechanisms and whose concentrations vary along time according to different factors. For instance, physiological adaptation resulting from various environmental conditions can be highlighted. More than segregated approach, integrated multi-omics approach would be beneficial for an efficient comprehension of communities' molecular biology (Hettich et al., 2013; Muller et al., 2013; Zhang et al., 2010).

2.1.1.2.1 Metatransciptomics

Metatransciptomics provide information about functionalities involved in the community through global transcriptome analysis. All ARNm from a community at a given moment are highlighted by metatransciptomics analysis. Gene translation is qualitatively (total genomic information is not necessary translated) and quantitatively (number of simultaneous transcript of a gene can vary) modulated according to different factors. However, due to the very short lifetime of ARNm molecules, such approach is not always easy to implement. Transcriptomics techniques are based either on direct analysis of ARNm sequences either on sequencing of complementary cDNA.

2.1.1.2.2 Metaproteomics

Metaproteomics provides a global view of the functional gene products in a community at a given moment. Contrary to metagenomics which provide functional potentialities, metaproteomics provides information about real functional activity. Metabolic activities implemented by community at a given moment can be established. Moreover, elucidating

metaproteome should also provide primordial information about interaction (competition as well as cooperation) occurring between the different actors of a community (Hettich et al., 2012).

Success of a metaproteome measurement relies on three factors: unbiased protein extraction from complex environmental matrix, efficient peptide/protein fractionation before detection and subsequent high-throughput peptide/protein identification (Hettich et al., 2012).

Separation and identification of proteins rely on two major strategies: either (1) gel electrophoresis based approach (i.e. 2D-GE and DIGE) which consists in separating proteins according to their isoelectric point and mass, and further MS identification; or (2) gel free methods which rely on coupling multidimensional liquid chromatography separation to automated tandem MS (LC-MS/MS). Compare to gel based methods, latter enable analysis of more complex proteomes so that they are more suitable for complex metaproteome analysis.

Hettich et al. (2013) and Lacerda and Reardon (2009) review several metaproteome measurements applied to environmental microbiology and biotechnology.

2.1.1.2.3 Metabolomics

Metabolomics aims for identifying and quantifying metabolites in biological system at a given time. When applied for characterizing overall metabolome from a community, "community metabolomics" term is proposed (Jones et al., 2014). Metabolites are products of metabolic pathway and final output of biological functions. They translate the activities of proteins and so determined final phenotypic properties (Washio et al., 2010). As proteins, metabolites can influence previous steps of cellular mechanism (metabolic regulation)

(Takahashi et al., 2012; Zhang et al., 2010) and their concentrations vary with physiological changes. Moreover metabolites directly affect cellular environment (e.g. increase of acidity).

Jones et al. (2014) applied such approach to soil communities living in contaminated sites. They apply [1]H Nucleic Magnetic Resonance to assess the metabolic profiles of the community after modified methanol-chloroform-water extraction (focus on aqueous-phase metabolites). Otherwise, chromatographic techniques coupled to mass spectrometry (GC-MS, LC-MS, CE-MS) are implemented. Moreover, such as interactomics for proteins interactions, specific omics field is dedicated to studies of metabolic fluxes. More precisely, fluxomics is the term associated to dynamic metabolic fluxes modelling which can be useful for investigation of complex large-scale metabolic systems (Zhang et al., 2010).

2.1.1.2.4 Combined approach: cytomics coupled with further molecular "-omics" techniques

Cytomics consider complex microbial communities at a single-cell level through flow cytometry analyses. On one hand, it works to provide structural as well as functional features for each cell and on the other hand it works to elucidate single-cell contribution to the overall community state. However, despite wide diversity of staining techniques, flow cytometry is not able to lead to a full understanding of complex microbial mechanisms. Therefore, cytomics is usually combined with other molecular "-Omics" such as proteomics (Jehmlich et al., 2010). Subpopulations of cells sorted thanks to flow cytometry measurement of cellular parameters are submitted to further molecular analyses for a deep understanding of cellular mechanisms involved in this particular subpopulation. Combined approach allows bypassing limitations inherent to both methods.

However, if this strategy is well established for eukaryote cell, its application to bacterial population is still restrained by several parameters. Major drawback is the limiting size of bacteria and the resulting low amount of intracellular species (i.e. volume of bacteria is approximately a thousand-fold smaller compared with blood cells, meaning that the protein content is reduced to the same degree (Jehmlich et al., 2010)). This weak cellular content is an issue for molecular analyses carried out on sorted subpopulation; especially for transcriptome analysis due to low half-time of bacterial mRNA (i.e. more time is necessary for sorting sufficient cell content for transcriptome analysis) (Müller and Nebe-von-Caron, 2010). In the case of proteome analysis, development of gel-free techniques decrease the number of required sorted cells for efficient proteome analysis; 10^6 cell against 10^9 for traditional gel based techniques (Müller and Nebe-von-Caron, 2010). Another limitation of this combined approach is the necessity of adequate staining for sorting bacterial subcommunities. Indeed, contrary to eukaryote cell, intrinsic characteristic of bacteria is not sufficient for efficient population segregation.

Despite these drawbacks, Jehmlich et al. (2010) highlighted that a combined approach (flow cytometry sorting with further proteomics analysis) is suitable for a segregated characterization of the different members from a microbial community. They considered artificial mixed culture composed of 2 types of microbial strains for validating their method. By comparing proteomes from sorted subpopulations (thought to each corresponds to one of the strains) with proteomes of strains obtained by pure culture, they confirmed the efficiency of this combined approach for the individual study of the different members of a community. Müller et al. (2012) also highlighted the efficiency of a combined approach for improving the comprehension of communities' structure and functionalities. They underline two main advantages of a preliminary cell sorting step: (1) cell abundancy measured by flow

cytometry provides quantitative information, (2) focusing on subcommunities decreases background information from molecular analyses and increase their resolution. In their case, T-RFLP profiling of sorted subcommunities provides phylogenetic affiliations (based on 16S rDNA profiling and clone library previously constructed for similar samples) and specific metabolic capacities (based on the use of gene as functional markers) of the different members of the sorted subcommunities.

2.1.2 Case study: management of anaerobic digester communities
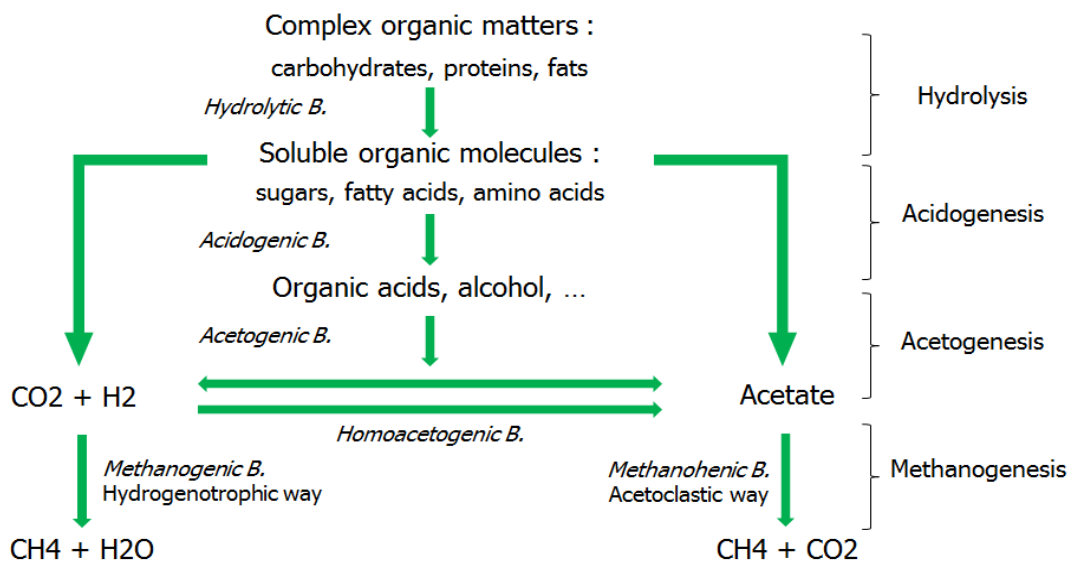
Figure 4 Anaerobic digestion metabolic pathways

Anaerobic digestion of organic matter is recognized as an attractive and commercially viable option as source of alternative energy. Methane can be obtained from the anaerobic digestion of different organic matters such as industrial residues, agricultural wastes, industrial wastewaters and municipal solid wastes. The anaerobic degradation process involves synergistic interactions among the various bacteria and methanogenic archea existing as a complex consortium. Four biochemical steps are involved: (1) hydrolysis of substrate polymers to monomers; (2) acidogenesis (i.e. conversion of monomers to volatile fatty acids, carbon dioxide, and hydrogen); (3) acetogenesis (i.e. production of acetate from

the intermediate metabolites); and finally (4) methanogenesis (i.e. conversion of acetate and carbon dioxide to methane).

2.1.2.1 *Who is there?*

As described above, anaerobic digestion involves complex microbial communities. Due to incomplete understanding of microbiome complexity, anaerobic digesters management is mainly based on process parameters and empirical knowledge rather than on microbial parameters. Improving the comprehension of the structure and functionalities of microbial communities would give a better perspective for the optimization of processes as it would allow an efficient control/steering of microbial population. Main required information to reach this goal, and notably to establish microbial based performance parameters, deal with the diversity, the evenness and the dynamics over time of microbial community (Carballa et al., 2015) (see section 2.1.1.1 for a description of MRM parameters).

Typically, molecular fingerprinting techniques are applied for the monitoring of microbial communities in bioreactors. de Araújo and Schneider (2008) confirm the reliability of sequencing segregated DNA bands from DGGE patterns for the identification of the major actors from a bioreactor community whereas Carballa et al. (2010) and Pycke et al. (2011) confirm suitability of the technique for the monitoring of population's dynamics. Particularly, Carballa et al. (2010) underline the highly dynamic behavior of communities from thermophilic and mesophilic digester for both *Bacteria* and *Archaea* (rate of change between 30% and 75% per 18 days which correspond to hydraulic retention time) even in stable performing periods. The DGGE analyses also provide information about communities' richness. As well in thermophilic as in mesophilic conditions, richness is higher for bacterial population than for archaeal population. Moreover, mesophilic conditions induce higher

diversity than thermophilic. Finally, the results of the molecular fingerprinting allow determining community organization. In both temperature conditions, a higher evenness is measured for bacteria than archaea. Furthermore, mesophilic conditions result in communities displaying higher evenness than communities shaped by thermophilic conditions for both bacteria and archaea. Interestingly, despite variability in term of absolute values, similar trends are depicted by different molecular fingerprinting techniques involving the three parameters.

However, over last years, with drawbacks inherent to molecular fingerprinting techniques in mind (exposed in previous sections), next generation sequencing methods have been applied by several authors in order to deeply characterize microbial population from anaerobic digester. Sundberg et al. (2013) highlight major bacterial and archaeal phylum involved in 21 anaerobic digesters fed with different type of solid waste. *Proteobacteria*, *Firmicutes*, *Bacteroidetes* and *Chloroflexi* were the four major phyla in the bacterial domain whereas *Methanosaeta* was shown to be dominant in term of archaeal genus diversity. Moreover, it has been shown that populations from 21 studied bioreactors are mostly composed of microorganisms from the *Clostridia* bacterial class (Sundberg et al., 2013). This suggests this type of microorganisms play a key role during anaerobic digestion. Especially, Kinet et al. (2015) highlight this type of microorganisms are key player for cellulose anaerobic digestion. The study led by Sundberg et al. (2013) also underlines the relationship existing between communities composition and (1) the nature of the substrate; and (2) the process temperature. These correlations are respectively confirmed by Ziganshin et al. (2013) and by Tian et al. (2015). Substrate nature and operational conditions (temperature and organic loading rates) dependencies are especially true for bacteria, whereas archaeal population is more affected by reactor environment (VFA concentrations, ammonium

concentration or even reactor design) (Carballa et al., 2015). Razaviarani and Buchanan (2014) also implemented next generation sequencing method for elucidating the mechanisms underlying population dynamics in bioreactor so that a link between population dynamics and reactor performance can be established. Parameters such as pH, alkalinity and particularly VFA concentrations were proved to be major factors shaping microbial community dynamics.

Flow cytometric monitoring has recently been considered as an alternative for elucidating complex microbial communities. On the basis of flow cytometry fingerprinting (CyBar method, see section 2.1.1.1.2 for a description of the method), Koch et al. (2013) highlighted adaptations of subcommunities towards artificial disturbances during a continuous digesting process. Further DNA based analyses underline subcommunities' phylogenetic compositions remain pretty stable whereas significant changes affect cell abundancies. Moreover, functionalities were attributed to several specific subcommunities through the establishment of statistical correlation between abiotic parameters and fingerprinting data.

2.1.2.2 Who is doing what?

Correlating populations' composition and dynamics resulting from molecular and cytometric techniques with operational conditions and performance parameters allows for functional insights. Especially, functional redundancy is thought to be beneficial for insuring digester performances in case of unpredictable stressful conditions. However these hypotheses are only based on statistical observations and need to be confirmed by molecular based proofs. Meta-omics techniques are suitable to provide such information and to clarify interspecies relationships. Complementary to metagenome analysis, Zakrzewski et al. (2012) applied metatranscriptomics for highlighting active microbial fraction involved in a full-scale

anaerobic digestion process and for identifying the most abundant enzymatic activities. As expected, transcripts corresponding to enzymes involved in the four different steps of the anaerobic digestion (substrate hydrolysis, acidogenesis, acetogenesis and methanogenesis) were the most abundant. Interestingly, it was also evidenced that most abundant species are responsible of majority of the community activity. With the same idea of elucidating metabolic pathways involved in anaerobic digestion in mind, metaproteomics approach was carried out in several studies. Heyer et al. (2015) reviewed some of these studies and highlight improvement of bioprocesses efficiency enabled by a such approach. Especially, the metaproteome of a community acting in an anaerobic digester fed with cellulosic substrate was investigated by Lü et al. (2014) complementary to its metagenome. Results from metaproteomics and further metagenomics suggest that cellulose and hemicellulose hydrolysis and fermentation results from the cooperative action of *Caldicellulosiruptor spp.* and *Clostridium thermocellum* species. Indeed several proteins thought taking part in cellulose and hemicellulose metabolism were identified and linked to these species. More precisely, proteome analysis suggests both bacteria are involved in cellulose degradation whereas hemicellulose is only degraded by *Clostridium thermocellum*. Other remarkable result is the absence of enzymes involved in acetoclastic methanogenesis, suggesting other pathway for acetate metabolism such as syntrophic acetate oxidation. Contrary, hydrogenotrophic methanogen pathway is clearly evidenced and can be linked to *Methanothermobacter*. This is in opposition with the habitual postulate attributing most of the methane formation to the acetoclastic pathway (Cabezas et al., 2015).

2.1.2.3 *Microbial based disturbance indicators*

As mentioned previously, environmental disturbances deeply affect microbiome structure. High attention must be paid to this issue in order to prevent the decrease of anaerobic

digester performances. Unfortunately, weak environmental modifications cannot always be highlighted by classical techniques dealing with measurement of bulk parameters. Therefore, alternative indicator must be implemented. Based on this statement and as change in microbiome structure is indicative of evolving conditions, population (in)stability assessment is considered as a promising alternative for early indication of processes disturbance.

As flow cytometry has been proved to be powerful for highlighting subcommunities disturbances towards operational conditions evolution (Koch et al., 2013b) and as this technique complies with several criteria such as (1) no impact of sampling and data acquisition; (2) dense sampling analysis (time required for analysis inferior to microorganisms generation time); (3) relatively low cost; (4) ease of manipulation and data interpretation (weak operator dependency); Koch et al. (2014b) suggest flow cytometry as routine approach for digester microbiome monitoring and more precisely (in)stability measurement.

Moreover, Carballa et al. (2015) proposed some microbial-based warning indicators based either on bacterial or archaeal communities. The first one is based on the monitoring of bacterial population evenness; a decrease from intermediate evenness (consider as optimum based on studies reviewing by Carballa et al., 2015) is considered as warning signals for anaerobic digester dysfunction. The second one is based on phylogenetic analysis of archaeal population; specific decrease in *Methanosaeta* population and/or active methanogen population highlight potential malfunction of the system.

2.1.2.4 Shaping reactor microbiome

Once information about microbiome structure is available, adequate microbial-based strategy can be applied for improving anaerobic digestion efficiency. First of all, a judicious inoculum selection is primordial to avoid long adaptation period. Inoculum diversity and structure, elucidated thanks to different previous mentioned methods, must be in agreement with the different operational conditions (T°, type of substrate …). However, despite a rational inoculum choice, degradation of recalcitrant substrates can be problematic and thus bioaugmentation treatment with specific population can be required. Lignocellulosic biomass is a typical recalcitrant substrate whose degradation is recognized as a limiting step during anaerobic digestion (Adney et al., 1991; Vavilin et al., 2008). Addition of a highly active cellulosic population has been proved for enhancing performance of digester dealing with such substrates (Hu et al., 2016; Martin-Ryals et al., 2015; Peng et al., 2014). Assessment of the establishment of the exogenous population among endogenous community informs about the efficiency of the treatment. For instance, Scherer and Neumann (2013) monitored the impact of compost addition on the distribution of different hydrolytic and methanogenic phylum in anaerobic digester on the basis of FISH analyses. They highlighted no increase of microbial count but important change affecting population structure, especially hydrogenotrophic methanogens. However, they were unable to affirm that positive effect on biogas production induced by compost addition is only due the increase of hydrogenotrophic methanogens. A positive impact of the hydrolytic bacteria, potentially brought with compost but not detected by the method implemented, cannot be excluded.

2.2 Phenotypic heterogeneity

Despite an homogeneous DNA content, isogenic bacterial populations display phenotypic heterogeneities during bioprocesses (Müller et al., 2010). Stochasticity affecting cellular biochemical reactions and epigenetics mechanisms are considered as the major contributors to phenotypic heterogeneity in bioprocess conditions (Delvigne et al., 2014). Cell cycle and cellular division induce heterogeneities at lower extent. Contrary to mutations affecting genetic information and only becoming an issue on timescales longer than typical timescales processes, all these traits are non-heritable and only have short term effects (Figure 5).
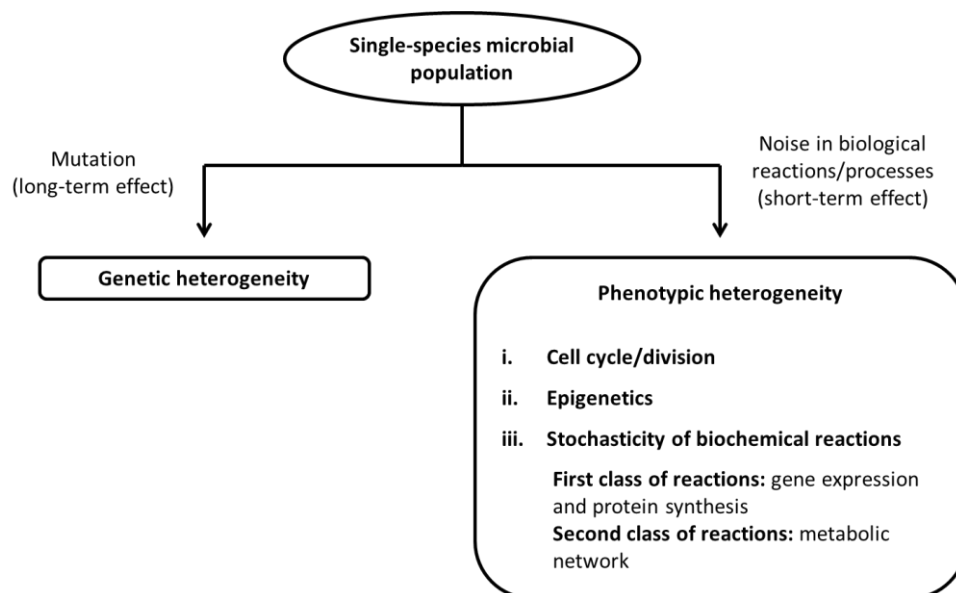

Figure 5 **Different sources of heterogeneity involved in microbial phenotypic diversification under bioprocessing conditions (Delvigne et al., 2014).**

As mentioned at Figure 5, stochasticity impacts two types of biochemical reactions: (1) stochasticity affects genes expression and further protein synthesis and (2) stochasticity affects metabolic network. This stochasticity results as well from internal noise as from external noise. Moreover, intrinsic and extrinsic noise can be distinguished among internal noise. Difference between these two kinds of noise is more easily comprehensive through their experimental measurement. Experimentations are led with two forms of GFP gene

reporter under the control of promoters regulated by same repressor. The two GFP present different emission wavelengths and are quantified thanks to fluorescence intensity. Intrinsic noise induces, in a same cell, unequal fluorescent intensity at respective wavelength mirroring differentiated gene expression. While extrinsic noise similarly affects both reporter genes in a same cell but induces inter-cellular differences. Variable content of proteins (influencing GFP gene expression) are responsible of these differences (Patnaik, 2006). In respect to experimental data obtained with *Escherichia coli*, *Bacillus subtilis* and *Saccharomyces cerevisiae*, among the two type of intra-cellular noises, extrinsic noise is majoritarian contributor to gene expression stochasticity (Patnaik, 2006).

As metabolic networks actors are modulated by gene expression, variability affecting this process results in heterogeneous metabolic activities. However, stochasticity affecting metabolic reactions cannot be only attributed to the stochastic nature of gene expression. ATP imbalance and molecular crowding can also be advanced as responsible for metabolic pathways heterogeneities (Delvigne et al., 2014). Furthermore, it is important to note that interactions between these two types of reactions are reciprocal. Indeed, if stochasticity in gene expression affects metabolic network through heterogeneous expression of metabolic enzymes, metabolic reactions in turn impact gene expression as some metabolites are key players of gene expression regulatory networks (de Lorenzo, 2014).

Finally, in large-scale bioprocess conditions, inherent heterogeneous micro-environmental conditions (e.g. emergence of temperature gradients, spatial heterogeneities in substrate distribution) reinforced stochasticity of cellular biochemical reactions. Fluctuating environment induce specific response at different levels of microbial physiology (transcriptional, translational and metabolic) (Delvigne et al., 2014). Particularly, chemical

heterogeneities such as concentration gradients of substrate, metabolites or dissolved oxygen take great part in physiological heterogeneities displayed by biofilm communities (Stewart and Franklin, 2008). Cells respond adaptively to local environment by turning on or off certain genes. Therefore, a wide range of distinct localized adaptations results from heterogeneous microenvironmental conditions in biofilm.

2.2.1 Impact of Phenotypic heterogeneities – Bet-hedging and division of labor strategies

Typically, phenotypic heterogeneities are recognized for negatively impacting global productivity of a process through the emergence of subpopulation exhibiting 'non-producer' or 'low-producer' phenotype (Figure 6) (Delvigne et al., 2014). However, phenotypical diversification, through bet-hedging and division of labor strategies, can also be beneficial for process robustness.
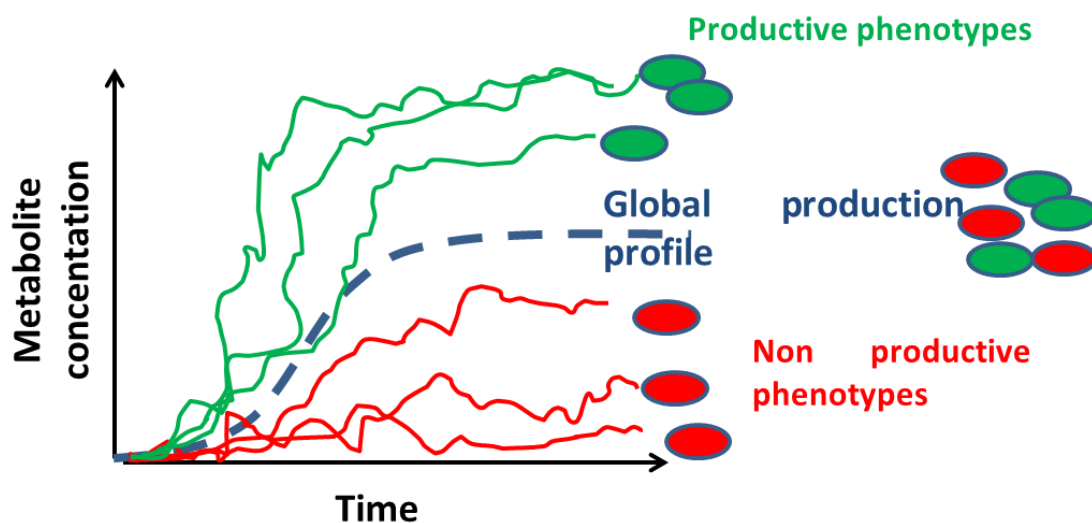

Figure 6 The accepted picture of the negative impact of microbial phenotypic heterogeneity on process performance. In the scheme a subpopulation with reduced protein synthesis capacity is shown that lowers the global yield of the bioprocess (F. Delvigne et al., 2014).

2.2.1.1 Bet-hedging strategy

In the environment, microbial populations exploit phenotypical diversification for facing adverse and fluctuating conditions. Thanks to random switch between phenotypes appropriate to each environment, cells are able to overcome more easily an environmental change (Martins and Locke, 2015). Such survival strategy, exploiting phenotypic variation, is proposed by some authors as bet hedging strategy (Martins and Locke, 2015; Veening et al., 2008). Among several co-existing phenotypes, only one is adapted to the environmental conditions encountered at a particular time so that long term population survival is ensured (figure 7).
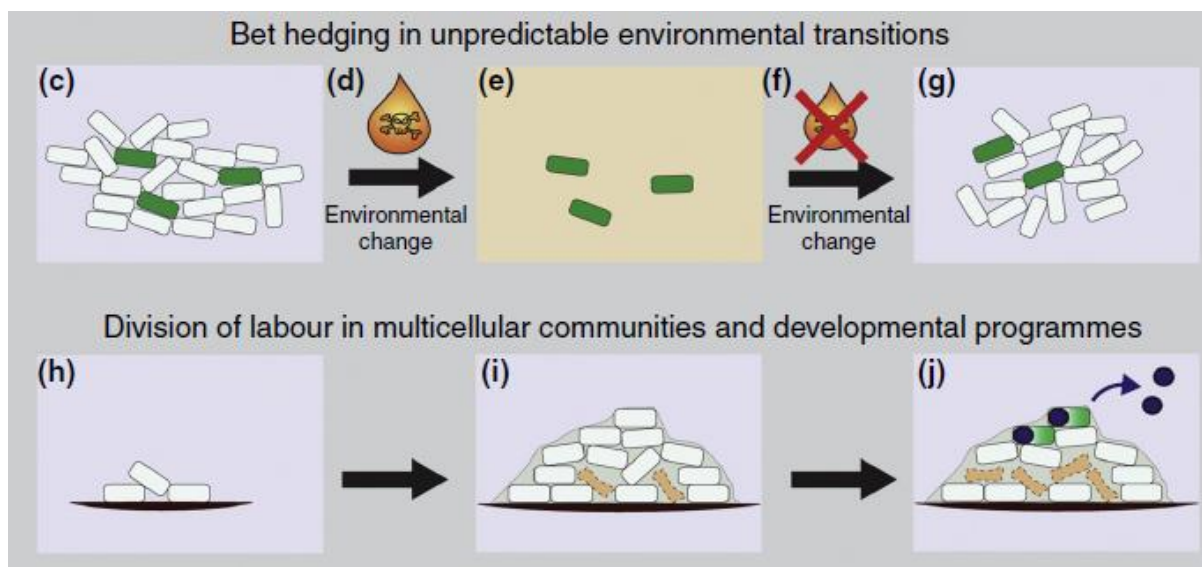

Figure 7 Cell-to-cell heterogeneity often reflects collective strategies, so one must 'zoom out' back to the level of populations to understand it. (c–g) The strategy of bet hedging copes with unpredictable environmental change. Often, a fraction of bacterial cells growing in rich media displays a phenotype that is not adequate for that particular environment (dark green cells) (c). These cells can, however, survive an unpredictable stress (e.g., antibiotic exposure) (d,e), thus allowing the population to survive and thrive again in the future (f,g). (h–j) Many microbes form multicellular aggregates and implement a strategy of division of labour, which allows the population to endure stress and activate developmental programmes. In microbial biofilms, the colony grows from a small aggregate (h) to a large sized community. Biofilms accommodate significant cell-to-cell heterogeneity. The growth of the structure relies on spatial and temporal regulation of apoptotic programmes (orange cells with dashed lines) (i), while the survival of the colony (j) is dependent on the successful sporulation, dispersal and future germination of a sub fraction of cells (dark blue spores) (Martins and Locke, 2015).

2.2.1.2 Division of labor

Phenotypic heterogeneities observed in an isogenic population can also result from a division of labor strategy. Especially, in single species biofilm communities, bacteria exploit stochastic gene expression (cells switching randomly between discrete physiological states by differentiating gene expression) in order to divide labor required for biofilm formation (Figure 7). Specialized subpopulations have been demonstrated in *B. subtilis* biofilm. Only a small proportion of cells has been evidenced for expressing gene encoding matrix components (Chai et al., 2008). The energy cost, inherent to the synthesis of these molecules, is saved by non-producing cells which benefit from matrix components produced by others. Similarly, biofilm maintenance and community survival/dispersion is ensured through functional specialization in mature communities (Figure 7). A part of the community multiply while others differentiate into several phenotypes such as matrix producer, spore former or competent cell through complex regulatory pathways (Stewart and Franklin, 2008). More than co-existence of specialized groups, division of labor strategy imply positive interactions between different groups. A colony formed of multiple cell types is assumed preforming better than a colony that consists solely of any one of them. Cooperative interactions of specialized sub-populations yield ecological benefit for the community. Therefore, similarly to bet-hedging strategy, division of labor increases population robustness and allows community to overcome environmental stresses.

2.2.2 Harnessing phenotypic heterogeneities in biofilm bioreactor

Recently, potential of biofilm has been pointed out for fine chemical/biological synthesis (Cheng et al., 2010). Division of labor (see section 2.2.1.2. for details) occurring in monospecies biofilm improves the robustness and the productivity of processes when comparing with planktonic cell strategy. In a process optimization perspective, such biofilm

microbiota cannot be considered as a 'blackbox'. The behavior of these "multi-phenotypes" multicellular communities should be deeply understood for the efficient management of biofilm reactors. Particularly, physiological heterogeneities and related underlying molecular mechanisms involved in biofilm formation and maintenance must be deeply understood for process optimization. Moreover, spatial heterogeneities should be taken into account.

2.2.2.1 Highlighting physiologic heterogeneities in single species biofilm

Stewart and Franklin (2008) report advantages and limitations of several experimental techniques which allow mapping physiological activities in biofilm. Among these ones, the most popular approach for characterizing localized levels of gene expression and resulting heterogeneities is probably the use of reporter genes fused to promoter regions of interest (Stewart and Franklin, 2008). Garcia-Betancur et al. (2012) applied this approach in order to analyze subpopulations responsible for matrix and surfactin production among *B. subtilis* biofilms. Fluorescent reporters are inserted into chromosome under the control of promoters related to the expression of the genes of interest; resulting fluorescence is then monitored either by fluorescence microscopy or flow cytometry. Such approach provides spatial and/or quantitative information about subpopulations expressing investigated genes. Moreover, coordination between cell differentiation pathways can also be highlighted through cells incorporating several reporters. It was notably shown for differentiation of *B. subtilis* biofilm subpopulations into matrix producers and cannibal phenotypes (López et al., 2009).

Fluorescent stains are also useful for discriminating differentiated subpopulations based on their physiological features. Used in combination, Syto9 (a permeable membrane stain which should penetrate and stain all cells) and propidium iodide (a dye which only stains cells

displaying compromised membrane) allows determining cell-membrane permeability. In this way, distinct resistance of subpopulations against antimicrobial treatment can be highlighted among biofilm subpopulatitons. Similarly, sporulating population can be evidenced thanks to Acridine Orange (AO) staining. Florescence intensity resulting from stain binding differs according to dye is bind to spore or vegetative cells.

Moreover, these techniques that allow evidencing different specialized subpopulations can be a basis for a cell sorting approach. Using selective GFP labelling of active and dormant subpopulations and cell sorting, Williamson and co-workers (2012) highlight differentiated tolerance of subpopulations face to antibiotics according to their growth state. Moreover, the different level of resistance can be related to the spatial localization of the subpopulation in the biofilm as this position has been evidenced for impacting their growth behavior. Cells located deeper in the biofilm matrix, displaying a viable but slow-growth state, has been proved as antibiotic-tolerant contrary to actively dividing cells from the top of the biofilm. As for multispecies microbial communities, subsequent physiological and/or molecular analyses (e.g. transcriptome or proteome analysis could be considered) can provide specific information about sorted subpopulations so that further management of the subpopulations is facilitate. By promoting given subpopulations and controlling/steering microbial heterogeneities in biofilm, overall efficiency of processes could be improved (MRM concept (Verstraete et al., 2007)).

2.3   Concluding remarks

The recent advances in culture-independent techniques, particularly in the field of "-Omics" based and flow cytometry technologies, allow for identifying the different actors of heterogeneous microbial communities. Sequencing of the gene coding for 16s rRNA allows

the identification of the different type of bacteria involved in multispecies communities while flow cytometry is proved for identifying the different phenotype present in an isogenic community. Given to the fast dynamics of microbial communities, advanced monitoring is required for ensuring efficient control/steering of bioprocess. Flow cytometry, fits well with such advanced analysis of the communities' structure whereas molecular techniques are not compatible with a routine implementation due to its cost and time requirement.

On another hand, there are still difficulties for explaining the role of each phenotype/genotype among multicellular communities. "Who is doing what?" key question is still only partially answered. Approaches combining the advantages of both flow cytometry and "-omics" techniques aims to elucidate community behavior at single-cell level. Nevertheless, despite recent development, several factors are still restricting for the implementation of such approach to bacterial communities. Particularly, weak concentration of intracellular molecular species targeted by "-omics" techniques in bacteria limits their efficient analysis (e.g. metabolites concentration in *E. coli* is equal to $\sim 3 \times 10^8$ copies/cell (http://bionumbers.hms.harvard.edu)).

2.4 References

Adney, W.S., Rivard, C.J., Shiang, M., Himmel, M.E., 1991. Anaerobic digestion of lignocellulosic biomass and wastes. Appl. Biochem. Biotechnol. 30, 165–183. doi:10.1007/BF02921684

Bent, S.J., Pierson, J.D., Forney, L.J., 2007. Measuring Species Richness Based on Microbial Community Fingerprints: the Emperor Has No Clothes. Appl. Environ. Microbiol. 73, 2399–2401. doi:10.1128/AEM.02383-06

Bombach, P., Hübschmann, T., Fetzer, I., Kleinsteuber, S., Geyer, R., Harms, H., Müller, S., 2011. Resolution of natural microbial community dynamics by community fingerprinting, flow cytometry, and trend interpretation analysis. Adv. Biochem. Eng. Biotechnol. 124, 151–181. doi:10.1007/10_2010_82

Cabezas, A., Araujo, J.C. de, Callejas, C., Galès, A., Hamelin, J., Marone, A., Sousa, D.Z., Trably, E., Etchebehere, C., 2015. How to use molecular biology tools for the study of the anaerobic digestion process? Rev. Environ. Sci. Biotechnol. 14, 555–593. doi:10.1007/s11157-015-9380-8

Carballa, M., Regueiro, L., Lema, J.M., 2015. Microbial management of anaerobic digestion: exploiting the microbiome-functionality nexus. Curr. Opin. Biotechnol. 33, 103–111. doi:10.1016/j.copbio.2015.01.008

Carballa, M., Smits, M., Etchebehere, C., Boon, N., Verstraete, W., 2010. Correlations between molecular and operational parameters in continuous lab-scale anaerobic reactors. Appl. Microbiol. Biotechnol. 89, 303–314. doi:10.1007/s00253-010-2858-y

Chai, Y., Chu, F., Kolter, R., Losick, R., 2008. Bistability and biofilm formation in Bacillus subtilis. Mol. Microbiol. 67, 254–263. doi:10.1111/j.1365-2958.2007.06040.x

Chao, A., 1984. Nonparametric Estimation of the Number of Classes in a Population. Scand. J. Stat. 11, 265–270.

Chao, A., Lee, S.-M., 1992. Estimating the Number of Classes via Sample Coverage. J. Am. Stat. Assoc. 87, 210–217. doi:10.1080/01621459.1992.10475194

Cheng, K.-C., Demirci, A., Catchmark, J.M., 2010. Advances in biofilm reactors for production of value-added products. Appl. Microbiol. Biotechnol. 87, 445–456. doi:10.1007/s00253-010-2622-3

de Araújo, J.C., Schneider, R.P., 2008. DGGE with genomic DNA: Suitable for detection of numerically important organisms but not for identification of the most abundant organisms. Water Res. 42, 5002–5010. doi:10.1016/j.watres.2008.09.011

de Lorenzo, V., 2014. From the selfish gene to selfish metabolism: Revisiting the central dogma. BioEssays 36, 226–235. doi:10.1002/bies.201300153

Delvigne, F., Zune, Q., Lara, A.R., Al-Soud, W., Sorensen, S.J., 2014. Metabolic variability in bioprocessing: implications of microbial phenotypic heterogeneity. Trends Biotechnol. 32, 608–616.

Di Bella, J.M., Bao, Y., Gloor, G.B., Burton, J.P., Reid, G., 2013. High throughput sequencing methods and analysis for microbiome research. J. Microbiol. Methods 95, 401–414. doi:10.1016/j.mimet.2013.08.011

Díaz, M., Herrero, M., García, L.A., Quirós, C., 2010. Application of flow cytometry to industrial microbial bioprocesses. Biochem. Eng. J., Invited Review Issue 2010 48, 385–407. doi:10.1016/j.bej.2009.07.013

Edenborn, S.L., Sexstone, A.J., 2007. DGGE fingerprinting of culturable soil bacterial communities complements culture-independent analyses. Soil Biol. Biochem. 39, 1570–1579. doi:10.1016/j.soilbio.2007.01.007

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27, 2194–2200. doi:10.1093/bioinformatics/btr381

Garcia-Betancur, J.C., Yepes, A., Schneider, J., Lopez, D., 2012. Single-cell analysis of Bacillus subtilis biofilms using fluorescence microscopy and flow cytometry. J. Vis. Exp. JoVE. doi:10.3791/3796

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A.J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J.Y.H., Zhang, J., 2004. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 5, R80. doi:10.1186/gb-2004-5-10-r80

Günther, S., Koch, C., Hübschmann, T., Röske, I., Müller, R.A., Bley, T., Harms, H., Müller, S., 2012. Correlation of Community Dynamics and Process Parameters As a Tool for the Prediction of the Stability of Wastewater Treatment. Environ. Sci. Technol. 46, 84–92. doi:10.1021/es2010682

Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J., Goodman, R.M., 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem. Biol. 5, R245–R249. doi:10.1016/S1074-5521(98)90108-9

Hettich, R.L., Sharma, R., Chourey, K., Giannone, R.J., 2012. Microbial metaproteomics: identifying the repertoire of proteins that microorganisms use to compete and cooperate in complex environmental communities. Curr. Opin. Microbiol., Ecology and industrial microbiology/Special section: Microbial proteomics 15, 373–380. doi:10.1016/j.mib.2012.04.008

Hettich, R.L., Pan, C., Chourey, K., Giannone, R.J., 2013. Metaproteomics: Harnessing the Power of High Performance Mass Spectrometry to Identify the Suite of Proteins That Control Metabolic Activities in Microbial Communities. Anal. Chem. 85, 4203–4214. doi:10.1021/ac303053e

Heyer, R., Kohrs, F., Reichl, U., Benndorf, D., 2015. Metaproteomics of complex microbial communities in biogas plants. Microb. Biotechnol. n/a-n/a. doi:10.1111/1751-7915.12276

Hu, Y., Hao, X., Wang, J., Cao, Y., 2016. Enhancing anaerobic digestion of lignocellulosic materials in excess sludge by bioaugmentation and pre-treatment. Waste Manag. 49, 55–63. doi:10.1016/j.wasman.2015.12.006

Jehmlich, N., Hübschmann, T., Salazar, M.G., Völker, U., Benndorf, D., Müller, S., Bergen, M. von, Schmidt, F., 2010. Advanced tool for characterization of microbial cultures by combining cytomics and proteomics. Appl. Microbiol. Biotechnol. 88, 575–584. doi:10.1007/s00253-010-2753-6

Jones, O.A.H., Sdepanian, S., Lofts, S., Svendsen, C., Spurgeon, D.J., Maguire, M.L., Griffin, J.L., 2014. Metabolomic analysis of soil communities can be used for pollution assessment. Environ. Toxicol. Chem. 33, 61–64. doi:10.1002/etc.2418

Kalyuzhnaya, M.G., Lidstrom, M.E., Chistoserdova, L., 2008. Real-time detection of actively metabolizing microbes by redox sensing as applied to methylotroph populations in Lake Washington. ISME J. 2, 696–706. doi:10.1038/ismej.2008.32

Kinet, R., Destain, J., Hiligsmann, S., Thonart, P., Delhalle, L., Taminiau, B., Daube, G., Delvigne, F., 2015. Thermophilic and cellulolytic consortium isolated from composting plants improves anaerobic digestion of cellulosic biomass: Toward a microbial resource management approach. Bioresour. Technol. 189, 138–144. doi:10.1016/j.biortech.2015.04.010

Kinet, R., Dzaomuho, P., Baert, J., Taminiau, B., Daube, G., Nezer, C., Brostaux, Y., Nguyen, F., Dumont, G., Thonart, P., Delvigne, F., 2016. Flow cytometry community fingerprinting and amplicon sequencing for the assessment of landfill leachate cellulolytic bioaugmentation. Bioresour. Technol. 214, 450–459. doi:10.1016/j.biortech.2016.04.131

Kleerebezem, R., van Loosdrecht, M.C.M., 2007. Mixed culture biotechnology for bioenergy production. Curr. Opin. Biotechnol. 18, 207–212. doi:10.1016/j.copbio.2007.05.001

Koch, C., Fetzer, I., Harms, H., Müller, S., 2013a. CHIC-an automated approach for the detection of dynamic variations in complex microbial communities. Cytom. Part J. Int. Soc. Anal. Cytol. 83, 561–567. doi:10.1002/cyto.a.22286

Koch, C., Fetzer, I., Schmidt, T., Harms, H., Müller, S., 2013b. Monitoring Functions in Managed Microbial Systems by Cytometric Bar Coding. Environ. Sci. Technol. 47, 1753–1760. doi:10.1021/es3041048

Koch, C., Gunther, S., Desta, A.F., Hubschmann, T., Muller, S., 2013. Cytometric fingerprinting for analyzing microbial intracommunity structure variation and identifying subcommunity function. Nat Protoc 8, 190–202. doi:10.1038/nprot.2012.149

Koch, C., Muller, S., Harms, H., Harnisch, F., 2014. Microbiomes in bioenergy production: from analysis to management. Curr Opin Biotechnol 27, 65–72. doi:10.1016/j.copbio.2013.11.006

Koch, C., Harnisch, F., Schröder, U., Müller, S., 2014a. Cytometric fingerprints: evaluation of new tools for analyzing microbial community dynamics. Front. Microbiol. 5, 273. doi:10.3389/fmicb.2014.00273

Lacerda, C.M.R., Reardon, K.F., 2009. Environmental proteomics: applications of proteome profiling in environmental microbiology and biotechnology. Brief. Funct. Genomic. Proteomic. 8, 75–87. doi:10.1093/bfgp/elp005

Li, K., Bihan, M., Yooseph, S., Methé, B.A., 2012. Analyses of the Microbial Diversity across the Human Microbiome. PLoS ONE 7, e32118. doi:10.1371/journal.pone.0032118

López, D., Vlamakis, H., Losick, R., Kolter, R., 2009. Cannibalism Enhances Biofilm Development in Bacillus subtilis. Mol. Microbiol. 74, 609–618. doi:10.1111/j.1365-2958.2009.06882.x

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.

Martin-Ryals, A., Schideman, L., Li, P., Wilkinson, H., Wagner, R., 2015. Improving anaerobic digestion of a cellulosic waste via routine bioaugmentation with cellulolytic microorganisms. Bioresour. Technol. 189, 62–70. doi:10.1016/j.biortech.2015.03.069

Martins, B.M., Locke, J.C., 2015. Microbial individuality: how single-cell heterogeneity enables population level strategies. Curr. Opin. Microbiol., Cell regulation 24, 104–112. doi:10.1016/j.mib.2015.01.003

Marzorati, M., Borin, S., Brusetti, L., Daffonchio, D., Marsilli, C., Carpani, G., Ferra, F. de, 2005. Response of 1,2-dichloroethane-adapted microbial communities to ex-situ biostimulation of polluted groundwater. Biodegradation 17, 41–56. doi:10.1007/s10532-005-9004-z

Marzorati, M., Wittebolle, L., Boon, N., Daffonchio, D., Verstraete, W., 2008. How to get more out of molecular fingerprints: practical tools for microbial ecology. Env. Microbiol 10, 1571–81. doi:10.1111/j.1462-2920.2008.01572.x

Muller, E.E.L., Glaab, E., May, P., Vlassis, N., Wilmes, P., 2013. Condensing the omics fog of microbial communities. Trends Microbiol. 21, 325–333. doi:10.1016/j.tim.2013.04.009

Müller, S., Harms, H., Bley, T., 2010. Origin and analysis of microbial population heterogeneity in bioprocesses. Curr Opin Biotechnol 21, 100–13. doi:10.1016/j.copbio.2010.01.002

Müller, S., Nebe-von-Caron, G., 2010. Functional single-cell analyses: flow cytometry and cell sorting of microbial populations and communities. FEMS Microbiol. Rev. 34, 554–587. doi:10.1111/j.1574-6976.2010.00214.x

Müller, S., Hübschmann, T., Kleinsteuber, S., Vogt, C., 2012. High resolution single cell analytics to follow microbial community dynamics in anaerobic ecosystems. Methods, Flow Cytometry and Cell Sorting: the Next Generation 57, 338–349. doi:10.1016/j.ymeth.2012.04.001

Patnaik, P.R., 2006. External, extrinsic and intrinsic noise in cellular systems: analogies and implications for protein synthesis. Biotechnol. Mol. Biol. Rev. 1, 121–127.

Peng, X., Börner, R.A., Nges, I.A., Liu, J., 2014. Impact of bioaugmentation on biochemical methane potential for wheat straw with addition of Clostridium cellulolyticum. Bioresour. Technol. 152, 567–571. doi:10.1016/j.biortech.2013.11.067

Postec, A., Urios, L., Lesongeur, F., Ollivier, B., Querellou, J., Godfroy, A., 2005. Continuous enrichment culture and molecular monitoring to investigate the microbial diversity of thermophiles inhabiting deep-sea hydrothermal ecosystems. Curr. Microbiol. 50, 138–144. doi:10.1007/s00284-004-4443-z

Pycke, B.F.G., 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. Water Sci. Technol. J. Int. Assoc. Water Pollut. Res. 63, 769–775. doi:10.2166/wst.2011.307

Ranjard, L., Poly, F., Nazaret, S., 2000. Monitoring complex bacterial communities using culture-independent molecular techniques: Application to soil environment. Res. Microbiol. 151, 167–177. doi:10.1016/S0923-2508(00)00136-4

Razaviarani, V., Buchanan, I.D., 2014. Reactor performance and microbial community dynamics during anaerobic co-digestion of municipal wastewater sludge with restaurant grease waste at steady state and overloading stages. Bioresour. Technol. 172, 232–240. doi:10.1016/j.biortech.2014.09.046

Read, S., Marzorati, M., Guimarães, B.M., Boon, N., 2011. Microbial Resource Management revisited: successful parameters and new concepts. Appl. Microbiol. Biotechnol. 90, 861–871. doi:10.1007/s00253-011-3223-5

Rogers, W.T., Holyst, H.A., 2009. FlowFP: A Bioconductor Package for Fingerprinting Flow Cytometric Data. Adv. Bioinforma. 2009, e193947. doi:10.1155/2009/193947

Rogers, W.T., Moser, A.R., Holyst, H.A., Bantly, A., Mohler III, E.R., Scangas, G., Moore, J.S., 2008. Cytometric fingerprinting: Quantitative characterization of multivariate distributions. Cytometry A 73, 430–441. doi:10.1002/cyto.a.20545

Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467.

Scherer, P., Neumann, L., 2013. "Methano-compost", a booster and restoring agent for thermophilic anaerobic digestion of energy crops. Biomass Bioenergy 56, 471–478. doi:10.1016/j.biombioe.2013.05.021

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., Weber, C.F., 2009. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Appl. Environ. Microbiol. 75, 7537–7541. doi:10.1128/aem.01541-09

Stewart, P.S., Franklin, M.J., 2008. Physiological heterogeneity in biofilms. Nat. Rev. Microbiol. 6, 199–210. doi:10.1038/nrmicro1838

Sundberg, C., Al-Soud, W.A., Larsson, M., Alm, E., Yekta, S.S., Svensson, B.H., Sørensen, S.J., Karlsson, A., 2013. 454 pyrosequencing analyses of bacterial and archaeal richness in 21 full-scale biogas digesters. FEMS Microbiol. Ecol. 85, 612–626. doi:10.1111/1574-6941.12148

Takahashi, N., Washio, J., Mayanagi, G., 2012. Metabolomic approach to oral biofilm characterization—A future direction of biofilm research. J. Oral Biosci. 54, 138–143. doi:10.1016/j.job.2012.02.005

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. doi:10.1016/j.watres.2014.11.001

Vavilin, V.A., Fernandez, B., Palatsi, J., Flotats, X., 2008. Hydrolysis kinetics in anaerobic degradation of particulate organic material: An overview. Waste Manag. 28, 939–951. doi:http://dx.doi.org/10.1016/j.wasman.2007.03.028

Veening, J.-W., Smits, W.K., Kuipers, O.P., 2008. Bistability, epigenetics, and bet-hedging in bacteria. Annu. Rev. Microbiol. 62, 193–210. doi:10.1146/annurev.micro.62.081307.163002

Verstraete, W., Wittebolle, L., Heylen, K., Vanparys, B., de Vos, P., van de Wiele, T., Boon, N., 2007. Microbial Resource Management: The Road To Go for Environmental Biotechnology. Eng. Life Sci. 7, 117–126. doi:10.1002/elsc.200620176

Washio, J., Mayanagi, G., Takahashi, N., 2010. Challenge to Metabolomics of Oral Biofilm: — From "What Are They?" to "What Are They Doing?"—. J. Oral Biosci. 52, 225–232. doi:10.1016/S1349-0079(10)80025-8

Williamson, K.S., Richards, L.A., Perez-Osorio, A.C., Pitts, B., McInnerney, K., Stewart, P.S., Franklin, M.J., 2012. Heterogeneity in Pseudomonas aeruginosa Biofilms Includes Expression of Ribosome Hibernation Factors in the Antibiotic-Tolerant Subpopulation and Hypoxia-Induced Stress Response in the Metabolically Active Population. J. Bacteriol. 194, 2062–2073. doi:10.1128/JB.00022-12

Wittebolle, L., Boon, N., Vanparys, B., Heylen, K., De Vos, P., Verstraete, W., 2005. Failure of the ammonia oxidation process in two pharmaceutical wastewater treatment plants is linked to shifts in the bacterial communities. J. Appl. Microbiol. 99, 997–1006. doi:10.1111/j.1365-2672.2005.02731.x

Wittebolle, L., Vervaeren, H., Verstraete, W., Boon, N., 2008. Quantifying Community Dynamics of Nitrifiers in Functionally Stable Reactors. Appl. Environ. Microbiol. 74, 286–293. doi:10.1128/AEM.01006-07

Zakrzewski, M., Goesmann, A., Jaenicke, S., Jünemann, S., Eikmeyer, F., Szczepanowski, R., Al-Soud, W.A., Sørensen, S., Pühler, A., Schlüter, A., 2012. Profiling of the metabolically active community from a production-scale biogas plant by means of high-throughput metatranscriptome sequencing. J. Biotechnol., Special Issue: ESF Conference (Microbes and Industrial Biotechnology) 158, 248–258. doi:10.1016/j.jbiotec.2012.01.020

Zhang, W., Li, F., Nie, L., 2010. Integrating multiple "omics" analysis for microbial biology: application and methodologies. Microbiology 156, 287–301. doi:10.1099/mic.0.034793-0

Ziganshin, A.M., Liebetrau, J., Pröter, J., Kleinsteuber, S., 2013. Microbial community structure and dynamics during anaerobic digestion of various agricultural waste materials. Appl. Microbiol. Biotechnol. 97, 5161–5174. doi:10.1007/s00253-013-4867-0

3 Scientific strategy and thesis structure

3.1 Scientific strategy

Anaerobic digestion is an interesting way for valorizing organic wastes such as ligno-cellulosic biomass. However, an optimization of the anaerobic digestion processes involving such substrates is still required. By these works, we try to bring some elements for improving the performances of these bioprocesses and more generally for improving performances of bioprocesses involving complex microbial communities.

In a first time, the negative impact of cellulose hydrolysis on the global performance of anaerobic digestion of (ligno-)cellulosic substrates is investigated. Particularly, an efficient mean of intervention for overcoming the limitation inherent to this step of the anaerobic digestion process is sought. Due to its relatively low cost and its simplicity of implementation, management of microbial resources involved in the bioprocess is an interesting option. Bioaugmentation treatment with cellulolytic bacterial strains adapted to environmental conditions encountered during anaerobic digestion is considered in this section for improving bioprocess performances.

In order to optimize the management of the microbial resources involved in bioprocceses, an efficient monitoring approach is required. Based on this statement, the second part of the works is devoted to the development of an efficient method for the monitoring of multispecies microbial communities during bioprocesses. Due to several limitations, classical culture-dependent techniques are non-adapted for this purpose. Therefore, culture-independent techniques must be taken into account. Particularly, displaying characteristics

required for a routine on-site implementation, flow cytometry is a promising option for the analysis of microbial communities. Therefore, the monitoring approach developed in this section integrates flow cytometry for the characterization of bacterial cells. Moreover, further bioinformatics approaches are integrated for obtaining operator independent fingerprint of the investigated communities. 16S rRNA gene sequencing is considered for validating data obtained through the designed flow cytometry based approach.

Many microbial communities involved in environmental biotechnologies form biofilm structure. Therefore, a monitoring technique adapted to such particular multicellular structures is required for elucidating and improving bioprocesses in which they are involved in. However, due to genotypic and phenotypic heterogeneity, functional monitoring of complex multispecies biofilm communities remain very tricky (Røder et al., 2016). In order to simplify the approach, our works will deal with monospecies biofilm communities. As for planktonic multispecies communities, efforts will be done for developing an efficient flow cytometry technique allowing fingerprinting of these "multi-phenotypes" communities.

3.2   Thesis structure

The body of the thesis which presents the more relevant results of the experimentations is divided into 3 chapters. The two first chapters address the management of complex multispecies communities while the third one deals with the monitoring of heterogeneities displayed by isogenic biofilm communities. Each chapter has been written in the format of a scientific publication.

In the first chapter, we describe the relevance of bioaugmentation treatment as an efficient way to improve the efficiency of cellulose anaerobic digestion. Moreover, the efficiency of 16S rRNA gene sequencing for describing complex communities during bioprocess is also depicted.

In the second chapter, we describe how flow cytometry fingerprinting approach can be considered as an efficient alternative to sequencing techniques for the routine monitoring of complex microbial communities in bioprocess conditions.

The third chapter focuses on the monitoring of metabolic heterogeneities among multicellular singlespecies biofilm communities. Particularly, the extension of the flow cytometry fingerprinting approach for the monitoring of metabolic heterogeneities during biofilm formation is considered in this chapter.

Finally, the last section is devoted to the general discussion of the results. Results from the three chapters are discussed globally and perspectives are proposed for overcoming the limitations encountered during the works.

3.3   References

1. Røder, H.L., Sørensen, S.J., Burmølle, M., 2016. Studying Bacterial Multispecies Biofilms: Where to Start? Trends Microbiol. 24, 503–513. doi:10.1016/j.tim.2016.02.019

4   Results

<div style="text-align:center">

CHAPTER I:

IMPROVING ANAEROBIC DIGESTION OF CELLULOSIC BIOMASS: TOWARD A MICROBIAL RESOURCE MANAGEMENT APPROACH

</div>

This chapter corresponds to the article entitled "*Thermophilic and cellulolytic consortium isolated from composting plants improves anaerobic digestion of cellulosic biomass: toward a microbial resource management approach*" (**Kinet Romain**, J. Destain, S. Hiligsmann, P. Thonart, L.Delhalle, B.Taminiau, G.Daube, F. Delvigne) published in **Bioresource Technology**, Volume 189, August 2015, Pages 138-144.

Abstract

A cellulolytic consortium was isolated from a composting plant in order to boost the initial hydrolysis step encountered in anaerobic digestion. Significant improvement of the cellulose degradation, as well as biogas production, was observed for the cultures inoculated with the exogenous consortium. 16S rRNA gene sequencing analyses pointed out a weak richness (related to the number of OTUs) of the exogenous consortium induced by the selective pressure (cellulose as sole carbon source) met during the initial isolation steps. Main microbial strains determined were strictly anaerobic and belong to the *Clostridia* class. During cellulose anaerobic degradation, pH drop induced a strong modification of the microbial population. Despite the fact that richness and evenness were very weak, the exogenous consortium was able to adapt and to maintain the cellulolytic degradation potential. This important result point out the fact that simplified microbial communities could be used in order to increase the robustness of mixed cultures involved in environmental biotechnology.

1 Introduction

Anaerobic digestion of organic matter is recognized as an attractive and commercially viable option as a source of alternative energy. Methane can be obtained by anaerobic digestion of different organic matters including industrial residues, agricultural wastes, industrial wastewaters and municipal solid wastes. The anaerobic degradation process involves synergistic interactions among various bacteria and methanogenic archea existing as a complex consortium. Four biochemical steps are involved, i.e. hydrolysis of substrate polymers to monomers, acidogenesis and conversion of monomers to volatile fatty acids, carbon dioxide, and hydrogen, acetogenesis and production of acetate from the intermediate metabolites, and finally methanogenesis with the conversion of acetate and carbon dioxide to methane. Lignocellulosic biomass represents an important part of the different reusable matter sources. This biomass is mainly composed of three different polymers i.e. cellulose, hemicellulose and lignin. Vavilin et al. (2008) highlighted the fact that the hydrolysis of lignocellulosic substrate is the rate-limiting step during biogas production from wastes rich in lignocellulosic plant fiber. Thus, effective hydrolysis of recalcitrant substrate is necessary for a profitable biogas production from fiber-rich wastes. Degradability improvement can be achieved by different physical (Pommier et al., 2010), chemical (Monlau et al., 2012) or biological (Zhang et al., 2011) pretreatments. However, pretreatments increase the cost of the global process and can generate various compounds interfering with microbial activities (Palmqvist & Hahn-Hägerdal, 2000). Therefore, a direct efficient bioconversion of raw lignocellulosic waste during anaerobic digestion process is more favorable.

As efficient lignocellulosic biomass degradation occurs in many kinds of natural environments by means of cooperative action of many microorganisms (cellulolytic and non-cellulolytic), management of consortia involved in these natural processes can be a source of sustainable solution. On the basis of this hypothesis, the "Microbial Resource Management" concept was developed by Verstraete et al. (2007). In order to efficiently manage endogenous microbial populations, three key questions must be answered, i.e. "Who is there?", "Who is doing what?", and "Who is doing what with whom?"(Verstraete et al., 2007). Marzorati et al. (2008) propose three parameters, based on molecular tools, to answer these questions: (i) the range-weighted richness (*Rr*) reflecting the carrying capacity of the system, (ii) the dynamics (*Dy*) reflecting the specific rate of species coming to significance, and (iii) functional organization (*Fo*), defined through a relation between the structure of a microbial community and its functionality (Marzorati et al., 2008).

Moreover, the integration of this concept to the enrichment culture techniques could be a powerful tool in order to engineer microbial consortia with given properties. Zuroff and Curtis (2012), reviewed some works about enriched cellulolytic natural consortia. However, only few works have emphasized on strict anaerobic enrichment. To date, most of the works have been focused on enrichment culture under aerobic or facultative anaerobic conditions. As well in aerobic conditions (Wongwilaiwalin et al., 2010) as in anaerobic conditions (Izquierdo et al., 2010) compost is a frequent microbial resource for enrichment culture. Nevertheless, other types of inocula have been tested for the development of efficient microbial cellulolytic communities, such as forest soil (Feng et al., 2011), or anaerobic digester sludge (Yan et al., 2012). Accordingly, based on enrichment method from compost samples, the present work has led to the design of an anaerobic thermophilic cellulolytic

microbial consortium able to degrade cellulose under conditions related to anaerobic digestion process.

2 Materials and methods

2.1 Biological Materials

Compost

Compost samples were collected from a composting center in Naninne, Belgium, exclusively supplied with green lignocellulosic waste. Sampling was realized during thermophilic phase of composting process at a temperature of 73.2°C.

Anaerobic sludge

Anaerobic sludge used as inoculum for BMP test originates from a full-scale anaerobic digester fed with agro-food organic waste and agricultural waste.

2.2 Cellulolytic microbial consortium preparation

BMP assay medium (Wang et al., 1994) supplemented with 10 g/l of cellulose (Whatman filter paper) is used as enrichment medium. Five grams of compost samples are inoculated in 125 ml bottles containing 45 ml of sterilized medium with a filter paper strip as an indicator for cellulase activity. Cultures are incubated at 37°C and 55°C under anaerobic conditions and without stirring. To establish anaerobic conditions, headspace of bottles, tightly capped with rubber septa and sealed with aluminum seals, is first flushed with carbon dioxide and after with oxygen-free nitrogen gas. Once the strip of filter paper begins to be degraded, 5 ml of culture are transferred into fresh enrichment medium. This process is repeated several times.

In parallel of this anaerobic enrichment, aerobic tests are led according to the procedure employed by Wongwilaiwalin et al. (2010). Five grams of compost samples are inoculated in flask containing 45 ml of PCS medium (0.1% yeast extract, 0.5% peptone, 0.5% CaCO$_3$, 0.5% NaCl, 1% filter paper, pH 8.0) with a filter paper strip as an indicator for cellulase activity. Cultures are incubated at 37°C and 55°C under aerobic conditions and without stirring.

2.3 Cellulose degradation capacity test

The cellulose degradation tests are led in BMP medium supplemented with 10 g/l of cellulosic material (filter paper and Avicell) for 7 days at 55°C under anaerobic conditions without stirring and with uninoculated medium as a control. At the end of the culture, centrifugation (8000 x g) allows separation of supernatant and pellet which includes residual substrate and microbial biomass. Pellet is then suspended in 100 ml acetic acid/nitric acid reagent (Feng et al., 2011) and heated at 100 °C for 30 min to remove the biological cells. Then, acetic acid/nitric acid treated suspension is filtered. The remaining cellulosic material is washed three times using 100 ml of distilled water each time. After washing and filtration, filtered solids are dried at 105 °C and determined gravimetrically (Feng et al., 2011).The weight loss of cellulosic materials is calculated by subtracting the weight of residual substrates from the total weight of cellulosic materials before degradation. Degradation ratio is calculated according to the equation:

$$Degradation\ ratio\ (\%) = \frac{M_t - M_r}{M_t} \times 100 \qquad (eq.1)$$

where M_t is total weight of the cellulosic materials before degradation and M_r is the weight of the residual substrates after degradation. All experiments were performed in triplicate and the average values were reported.

2.4 Anaerobic digestion test (Biochemical Methane Potential assay)

BMP assay medium (Wang et al., 1994) is used. Experiments are performed according to the procedure published by Wang et al. (1994) and all experiments are carried out in triplicate in 125 ml sterile glass serum bottles. 0.5 g of filter paper, as cellulosic substrate, or 0.5g of mechanically treated paper paste (composition: 53% of holocellulose, 32% of lignin and 15% of others compounds) , as lignocellulosic substrate, are introduced into bottles containing 45 ml of BMP medium, and 5 ml of inocula. Inocula consisted of either only anaerobic sludge or a mix 50:50 of anaerobic sludge and isolated cellulolytic consortium. pH is adjusted with a 0.5 M KOH solution to achieve an initial pH of 7.3 in each sample, and a maximum variation during the culture period of pH ±1 is maintained. The sample bottles are capped tightly with rubber septa and sealed with aluminum seals. Nitrogen is passed into the bottles to flush out air and other gases before incubation at 55°C.

The biogas productions and the composition of produced biogas are monitored for 90 days. Biogas production are regularly collected using a syringe fitted with a needle and the composition is monitored according to procedures described in section 2.7 (gas phase analysis). First, collected biogas (or a fraction of collected biogas) is injected in 100 ml gas replacement equipment containing a 9 M KOH solution. This KOH procedure allows for a fast determination of carbon dioxide concentration in gas phase. In a second time, CH_4 and H_2 concentration are determined. To do this, 25 µl of bottles' gas phase are analyzed by chromatography. During this second analysis, CO_2 concentration determined by KOH procedure is confirmed.

Blank samples are realized to highlight the impact of organic matter introduced with the two different inocula. Blank samples consist of 5 ml of inocula and 45 ml of BMP medium. No

energetic substrate is added to blank samples. Blank samples' biogas productions are subtracted from samples' production inoculated with the corresponding inocula.

2.5   Cellulose degradation kinetics

Independent bottles are prepared in triplicate for each sampling time. Growth conditions similar to 2.4 but without pH adjustment are used. For each sampling time, cellulose degradation ratio, metabolites accumulation and population composition are analyzed. Method presented in section 2.3 is used for the determination of cellulose degradation ratios. Volatile Fatty Acids (VFA), ethanol and glucose concentration in liquid phase are analyzed according to procedure described in section 2.7 (Liquid phase analysis). For each sampling time, bacteria identifications are realized on the sample presenting the highest degradation ratio and according to method developed by Delcenserie et al. (2014) described in section 2.6 (16S rDNA sequencing)

2.6   16S rDNA sequencing

Total DNA extraction

Total DNA is isolated from each primary suspension using the DNeasy Blood & Tissue DNA extraction kit (Qiagen Benelux B.V., Venlo, the Netherlands), following the manufacturer's extraction for gram-positive bacteria. Briefly, the bacteria present in the suspension are lysed for 1 hour at 37°C in a lysis buffer (Tris-HCL 20mM, sodium EDTA 2mM, triton x-100 1.2% and lysozyme 20mgml^{-1}. The suspension is then treated with proteinase K for 1 hour at 56°C. The resulting suspension is mixed with ethanol and loaded on a Dneasy Mini-spin column. The DNA is then washed 2 times with QIAgen buffers and eluated into DNase/RNase-free water and its concentration and purity are evaluated by optical density using the NanoDrop ND-1000 spectrophotometer (Isogen Life Science NV, Sint-Pieters-

Leeuw, Belgium). The DNA samples are stored at −20°C until use in 16S rDNA amplicon pyrosequencing analysis.

16S rDNA Gene Library Construction and Pyrosequencing

The 16S PCR libraries were generated for the 5 samples. The primers E9-29 and E514-530 (Brosius et al., 1981), specific to bacteria, were selected for their theoretical ability to generate the least bias of amplification capability among the various bacterial phyla (Wang & Qian, 2009). The oligonucleotide design included 454 Life Sciences A or B sequencing titanium adapters (Roche Diagnostics Belgium NV, Vilvoorde, Belgium) and multiplex identifiers fused to the 5' end of each primer. The amplification mix contained 5 U of FastStart highfidelity polymerase (Roche Diagnostics Belgium NV), 1× enzyme reaction buffer, 200 μM deoxynucleotide triphosphates (dNTP; Eurogentec SA, Liege, Belgium), 0.2 μM concentration of each primer, and 100 ng of genomic DNA in a volume of 100 μL. Thermocycling conditions consisted of a denaturation step at 94 °C for 15 min, followed by 25 cycles of 94 °C for 40 s, 56 °C for 40 s, 72 °C for 1 min, and a final elongation step of 7 min at 72 °C. These amplifications were performed on an EP Master system gradient apparatus (Eppendorf AG, Hamburg, Germany). The PCR products were run on a 1 % agarose electrophoresis gel and the DNA fragments were extracted and purified using an SV PCR purification kit (Promega Benelux B.V., Leiden, the Netherlands). The quality and quantity of the products were assessed using a PicoGreen double-stranded DNA (dsDNA) quantitation assay (Isogen Life Science NV). All libraries were run in the same titanium pyrosequencing reaction using Roche multiplex identifiers. All amplicons were sequenced using the Roche GS-Junior Genome Sequencer instrument (Roche Diagnostics Belgium NV).

16S rDNA Data Processing

The 16S rDNA sequence reads were processed using the MOTHUR software package (Schloss et al., 2009). The quality of all the sequence reads was assessed by using the PyroNoise algorithm implemented in MOTHUR and filtered according to the following criteria: minimal length of 425 bp, an exact match to the barcode, and 1 mismatch allowed to the proximal primer. The sequences were checked for the presence of chimeric amplifications using the UCHIME algorithm(Edgar et al., 2011). The resultant read sets were compared with a reference data set of aligned sequences of the corresponding region derived from the SILVA database of full-length rDNA sequences (http://www.arb-silva.de/) implemented in MOTHUR (Pruesse et al., 2007). The final reads were clustered into operational taxonomic units (**OTU**) with the nearest neighbor algorithm using MOTHUR with a 0.03 distance unit cutoff. A taxonomic identity was attributed to each OTU by comparison with the SILVA database (80 % homogeneity cutoff). As MOTHUR is not dedicated to taxonomic assignment beyond the genus level, all unique sequences for each OTU were compared with the SILVA data set (version 111), using the BLASTN algorithm (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome(Altschul et al., 1990)). For each OTU, a consensus detailed taxonomic identification was given based upon the identity (less than 1 % of mismatches with the aligned sequence) and the metadata associated with the most frequent hits (validated bacterial species or not).

BioSample Accession Numbers

All the BioSample sequences were deposited at the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov) and are available under the BioProject ID PRJNA273787

2.7   Analytical methods

Gas phase analysis

KOH operating procedure

Water containing 9 M KOH in 100 ml gas replacement equipment is used to monitor the carbon dioxide concentration in gas phase. The absorption potential of the KOH solution is regularly measured using gas mixtures containing 0, 20, 35, 80, and 100 % CO2.

Gas phase chromatography

Hewlett Packard 5890 Series II gas chromatograph (GC; Agilent Technologies, Santa Clara, CA, USA) equipped with a 30 m long, 0.32 mm id Alltech GAS PRO GSC column (Grace, Deerfield, IL, USA) in series with a 20 m long, 0.25 mm id Chrompack CARBOPLOT P7 column (Agilent Technologies) and a thermal conductivity detector allows H_2, CH_4, and CO_2 separation and detection. The carrier and reference gas for determining the proportions of CH_4 and CO_2 in the biogas is He. A mixture of N_2 (15 %), CO_2 (35 %), and CH_4 (50 %) is used as standard. For H_2 analysis, carrier and reference gas is N_2. A mixture of H_2 (80 %) and CO2 (20 %) is used as standard. The GC injection port, the thermal conductivity detector chamber, and the oven are maintained at 90, 110, and 55 °C, respectively.

As well for samples analysis as for standardization, 25 µl of gas are injected on-column.

Liquid phase analysis

Volatile fatty acid (VFA) concentrations in the culture medium were determined using an Agilent 1110 series high performance liquid chromatograph (HPLC; Agilent Technologies) equipped with a Supelcogel C-610H column (Sigma-Aldrich, St Louis, MO, USA) preceded by a Supelguard H precolumn (Sigma-Aldrich). The columns were kept at a temperature of 40 °C, and the isocratic mobile phase was 0.1 % H_3PO_4 (in ultrapure, ''milliQ'', water), at a flow rate of 0.5 ml min^{-1}. A differential refraction index detector, kept at 35 °C, was used. This analysis took 35 min at a maximum pressure of 60 bar.

2.8 Ecological parameters calculation

Inspired by resource management indices (Marzorati et al., 2008) and based on amplicon sequencing results, ecological parameters are calculated. Richness parameter corresponds to the number of observed OTUs at species level for each sample. Population evenness is expressed by the Shannon equitability (E_H) parameter calculated according to the following formula (eq.2):

$$E_H = H/H_{max} \quad \text{(eq.2)}$$

where *H* represents the Shannon's diversity index, calculated according to (eq.3) :

$$H = -\sum_{i=1}^{S} p_i \ln p_i \quad \text{(eq.3)}$$

where *S* represents the total number of observed species in the sample and p_i the proportion of *S* made up of the ith species. Equitability assumes a value between 0 and 1 with 1 being complete evenness.

The community dynamics is evaluated via two parameters: change (%) and rate of change (%.h^{-1}). % of change is calculated for each sampling time compared to inoculum (time = 0h) in accordance to the formula (eq.4):

$$\% \, change = 100 \, \% - similarity \,\,\, (eq.4)$$

where similarity values are based on the Pearson product-moment correlation (Marzorati et al., 2008). The rate of change expresses how rapidly the population evolves between two samples according to the formula (eq.5):

$$rate \, of \, change_{(x,y)} \, (\%.\, h^{-1}) = \frac{\% \, change_{(x,y)}}{h_y - h_x} \,\,\, (eq.5)$$

where *h* represents the number of hours between sampling and inoculation.

3 Results and discussion

3.1 Consortium isolation

Efficient thermophilic cellulose-degrading microbial community is obtained after successive subcultivations of compost, as microbial inocula, under specific conditions. The consortium is able, in seven days, at high temperature (55°C) and in anaerobic conditions with pH control and without stirring, to degrade filter paper extensively. 98.7 % ±0.3 of maximal weight loss, equivalent to a global degradation rate of 0.06 g.l^{-1}.h^{-1}, is obtained. In comparison, in the same conditions, consortia isolated by Izquierdo et al. (2010) display a global degradation rate of 0.05 g.l^{-1}.h^{-1}. The degradation capacity of the consortium is also similar to the one obtained by Tachaapaikoon et al. (2012) with pure cultures of *Clostridium thermocellum*, an anaerobic thermophilic species recognized for its high cellulose-degrading capacity. Filter

paper is not the only cellulosic substrate efficiently degraded by the consortium. Contrary to Viravaidya et al. (2011), high level of crystallinity has no negative impact on substrate degradation. Isolated microorganisms can degrade, with high efficiency and the same degradation rate (0.06 $g.l^{-1}.h^{-1}$), crystalline cellulose (98.4 % ±0.2 maximal weight loss). Isolated consortium seems to have a complete set of efficient cellulolytic enzymes to degrade different forms of cellulose.

Surprisingly, no growth was recorded in aerobic conditions. Consortium presents exclusive strict anaerobic characteristics. Based on compost samples were taken during aeration step and at low depth in the greenwaste mass, this property is unexpected. Previous consortia enriched from compost were, mostly, composed of aerobic and anaerobic species ((Wongwilaiwalin et al., 2010), (Feng et al., 2011)). Aerobes probably eliminating oxygen before anaerobic cellulolytic bacteria degrade substrate under anaerobic conditions.

The thermophilic nature of the consortium is more comprehensive as inocula come from compost with a recorded temperature of 73°C.

3.2 Impact of cellulolytic microbial consortium on biogas production

The impact of consortium on biogas production is evaluated under thermophilic conditions during biochemical methane potential tests (adapted method from Wang et al. (1994)). Experiments are led at 55°C under anaerobic conditions with pH adjustment on cellulosic and lignocellulosic substrates, respectively filter paper and mechanically treated paper past. For both inoculums (i.e. untreated anaerobic sludge and bioaugmented sludge), the values discussed below express the productions measured for the replicates displaying the highest final cumulative volume.

Figure 1 presents the evolution of cumulated volumes of biogas and methane generated from substrates during 90 days by industrial sludge alone and a mix 50:50 of this sludge and consortium. For all the experiments, H_2 concentrations are negligible. CO_2 and CH_4 are the sole components present in significant proportion in produced biogas. For the filter paper digestion, the results highlight a positive effect of the isolated consortium on biogas and methane production all along the process (90 days). Mix inoculum induces higher biogas and methane production than digester sludge (table 1). These increases represent gains of 14% for biogas production and 15% for accumulated methane. The addition of consortium allows increasing bioconversion rates of cellulosic substrate into biogas. By comparison with the theoretical yield of biogas production from carbohydrates (746 ml.g^{-1}, calculated according to Buswell equation (Buswell & Mueller, 1952)), our experiments point out bioconversion rates of 85% and 97%, respectively when the industrial sludge is used alone and when the consortium is added. The addition of an efficient cellulolytic population offsets the lack of efficiency of industrial sludge.

Consortium addition also provides positive effect during lignocellulosic substrate anaerobic digestion. Total biogas and methane production increase (table 1). After the addition of the cellulolytic population, respective gains of 15% and 12% are obtained. Despite positive impact of the cellulolytic microorganisms, substrate structure still induces high recalcitrance to bioconversion. High lignin content of wood, unaffected by mechanical process applied to obtain paper paste, is probably responsible of this recalcitrance. Indeed, hardly degraded anaerobically, lignin protects cellulose fibers from anaerobic digestion. Accordingly, Salehian et al. (2013) highlighted untreated wood substrate are not easily converted to methane.
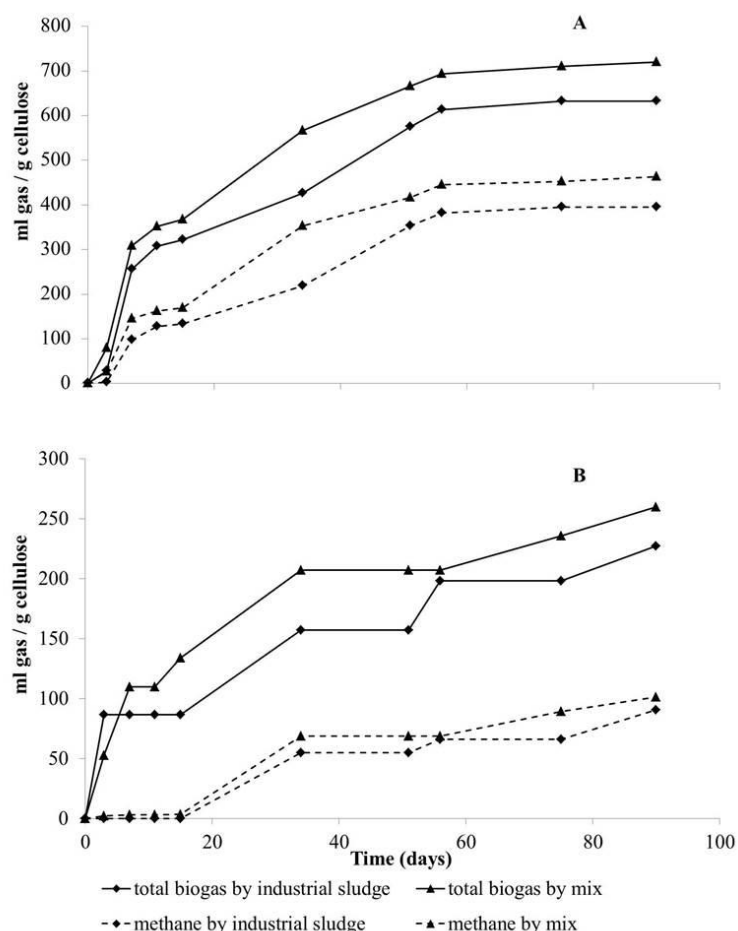
Figure 1 Evolution of total biogas and methane production (ml.g cellulose^{-1}) during anaerobic and thermophilic (55°C) digestion of (A) cellulosic substrate (10 g.l^{-1} filter paper) (B) lignocellulosic substrate (10 g.l^{-1} mechanically treated paper paste) by (1) industrial digester sludge (10% v/v) and (2) mix 50:50 of digester sludge and isolated consortium (10% v/v). For both inoculums, the values express the productions measured for the replicates displaying the highest final cumulative volume.

Table 1 Total biogas and methane production during anaerobic and thermophilic (55°C) digestion of cellulosic substrate (10 g.l^{-1} filter paper) and lignocellulosic substrate (10 g.l^{-1} mechanically treated paper paste) by industrial anaerobic digester sludge (10% v/v) and mix 50:50 of digester sludge and isolated consortium (10% v/v). Bioconversion rate of cellulosic substrate calculated according to Buswell equation. For both inoculums, the values express the productions measured for the replicates displaying the highest final cumulative volume.

	Cellulosic substrate		Lignocellulosic substrate	
	Anaerobic digester sludge	Digester sludge and consortium mix	Anaerobic digester sludge	Digester sludge and consortium mix
Biogas production (ml.g cellulose^{-1})	633	720	227	260
Methane production (ml.g cellulose^{-1})	395	463	90	101
Methane proportion (%)	62	64	40	39
Bioconversion rate (%)	85	97	-	-

3.3 Characterization of anaerobic cellulose degradation by a cellulolytic consortium isolated from composting plant

For the characterization of the anaerobic degradation of cellulose by the isolated cellulolytic consortium in thermophilic conditions, tests are performed at 55°C in minimal cellulosic medium with filter paper as only energetic source and without pH correction. Substrate degradation (figure 2), fermentation products (figure 3 A), pH (figure 3 B) and microbial community structure (Figure 4) are analyzed until stabilization.

Figure 2 shows the kinetic of filter paper degradation by the microbial community in anaerobic and thermophilic conditions. Degradation kinetics curve displays a typical sigmoidal growth curve shape, divided in three distinct phases. No degradation is observed during the first 19 hours, pointing out a possible adaptation mechanism of the consortium. Accordingly, no metabolites production is observed during this lag phase. A period of significant degradation follows this first period. Over a period of three days (between the 19th and the 91st hours of incubation), 40% of the total filter paper is degraded. Maximum cellulolytic potential of the microbial community is reached during this phase.
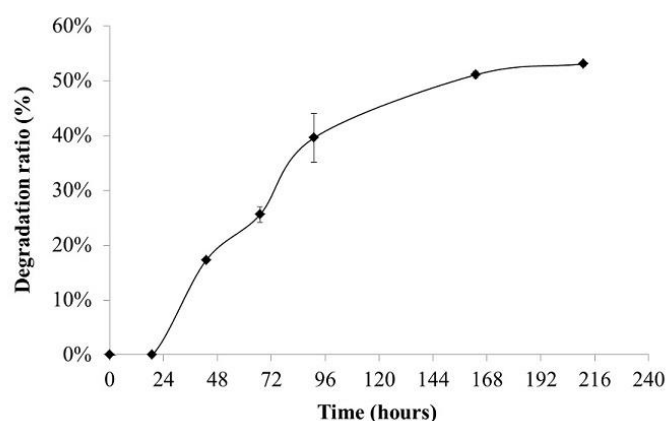

Figure 2 Filter paper (1% DM) degradation kinetics by isolated consortium (10% inoculum). Tests realized in triplicate at 55°C and in anaerobic conditions.
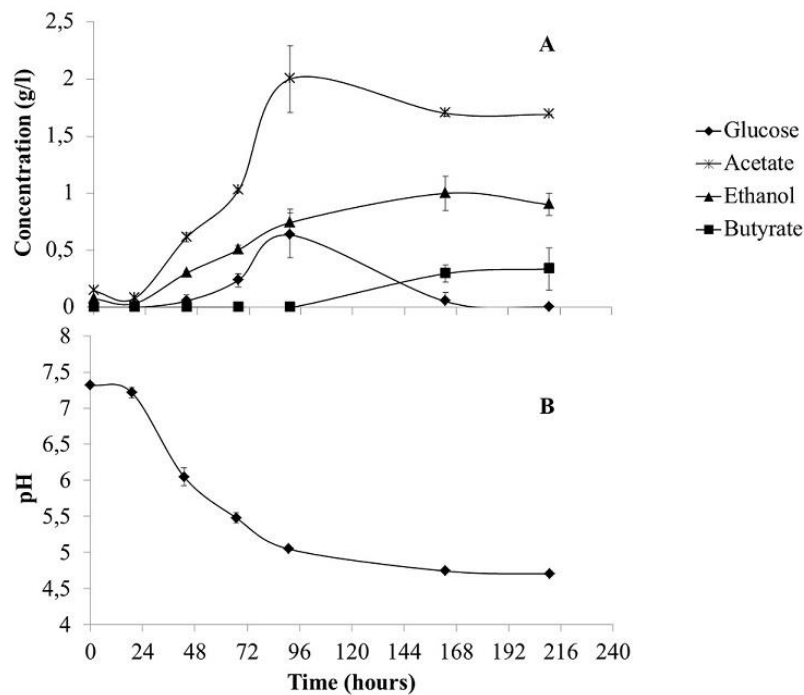
Figure 3 Evolution of (A) VFA's concentration (g.l^{-1}) and (B) pH during anaerobic thermophilic (55°C) cellulose (filter paper 10 g.l^{-1}) degradation by isolated consortium (10% v/v).

Between the 19th hour and the 67th hour of incubation, degradation ratio of 1.32 g.l^{-1}.h^{-1} ±0.02 is obtained. Furthermore, during this second phase, different metabolites are produced in significant quantity. Acetate is the main volatile fatty acid released. Ethanol and glucose are the other metabolites present in the liquid phase. The accumulation of glucose (0.63 g.l^{-1}), only component of cellulose, is weaker than acetate (2.21 g.l^{-1}) and ethanol (0.66 g.l^{-1}). Sugars released after cellulose degradation by the cellulolytic consortium are directly metabolized. This efficient cellulose degradation in neutral and weak acid pH growth conditions results from the synergistic metabolic activities exhibited by the different phylotypes present in the consortium. During the nineteen first hours, bacterium belonging to *Defluviitalea* genus is largely predominant with relative abundances of 74.5% and 70.2% respectively just after the inoculation and after nineteen hours of process. *Defluviitalea* genus is composed of only one species named *Defluviitalea saccharophila*. Strains belonging

70

to this genus was previously isolated by Jabari et al. (2012) and exhibited optimal growth parameters in an environment similar to our experimental conditions and was able to metabolize cellobiose, as well as acetate.
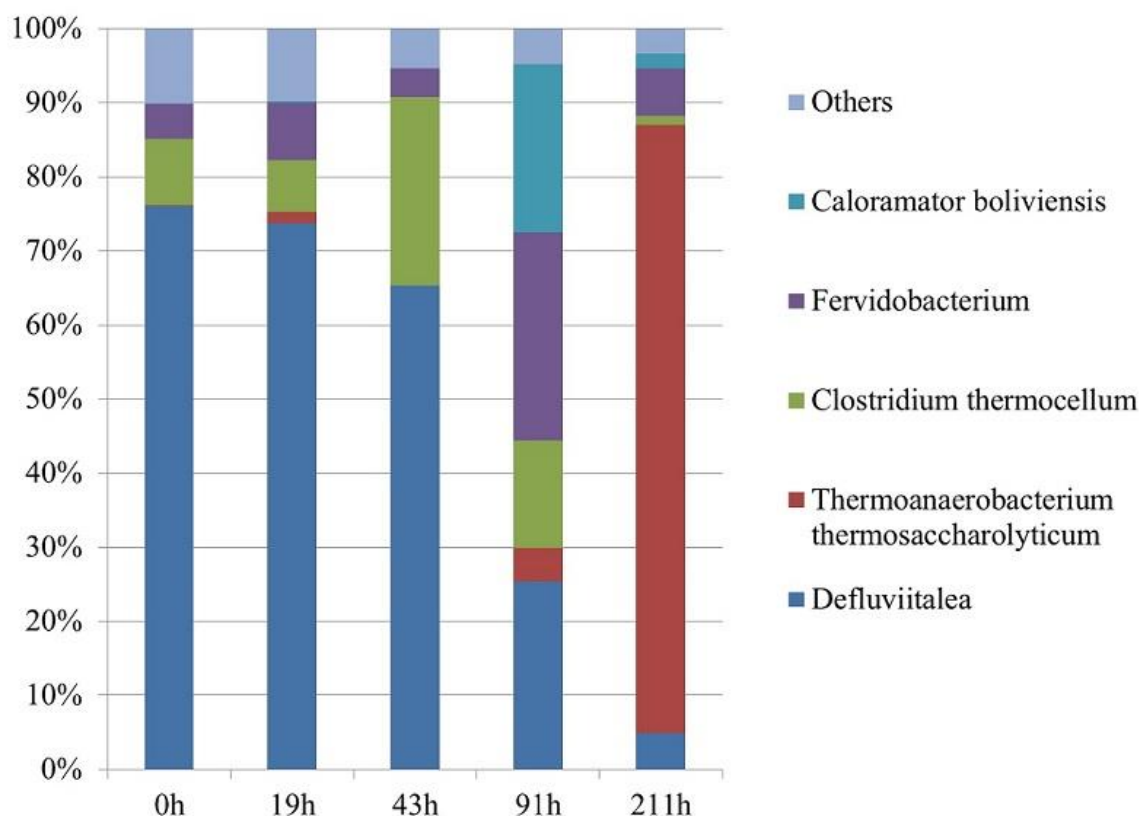

Figure 4 Population structure dynamic during anaerobic thermophilic (55°C) degradation of cellulose (filter paper 10 g.l^{-1}) by isolated consortium (10% v/v).

During this first step, *Clostridium thermocellum* and species belonging to *Fervidobacterium* genus are the two others main phylotypes (Figure 4). After 43 hours, the relative abundance of *Clostridium thermocellum* increases threefold (8.76% to 24.7%). This increase is proportional to the decrease observed for the *Defluviitalea* (74.5% to 63.5%). Whereas *Clostridium thermocellum* is very efficient for cellulose hydrolysis, this microbial strain cannot completely utilize the degradation products, cellobiose and glucose (Liu et al., 2008). However, the presence of other phylotypes, able to utilize these products, allows avoiding high accumulation and the resulting inhibition of cellulase activity. Furthermore, in

agreement with Liu et al. (2008), the presence of other species allows for the metabolization of lactate, one the main end-products of cellulose fermentation by *Clostridium thermocellum*. The accumulation of organic acids induces a significant decrease of the pH value during our cultivation tests (Figure 3B). At the end of this first phase of intensive cellulose digestion, a pH value of 5.05 is observed. During the last hours of this first period, acidic conditions favor the development of *Caloramator boliviensis* (21.6%), a thermophilic ethanol producing bacterium able to utilize (hemi-)cellulosic substrate (cellobiose, carboxymethylcellulose, xylan) (Crespo et al., 2012). During the second phase, acidification (pH under 5) leads to a global slowdown of the degradation process. Important shift in volatile fatty acid production and microbial population evolution are also recorded. Between the 91st and the 211th hours, the degradation process slowed-down (0.01 $g.l^{-1}.h^{-1}$) to finally attain 53% of degradation. Butyrate is released and attains a final concentration of 0.34 $g.l^{-1}$. Contrary to acetate, which is consumed (final concentration of 1.69 $g.l^{-1}$), ethanol concentration still rise during this period to attain final value of 0.90 $g.l^{-1}$. Acidic conditions promote the development of totally different species. Indeed, acidic conditions lead to the disappearance of almost all species except *Thermoanaerobacterium thermosaccharolyticum*, which is becoming largely predominant with 80.8% of the total population. *Thermoanaerobacterium thermosaccharolyticum* is also able to utilize cellulose and cellobiose with butyrate as one of the main metabolite (Ren et al., 2008).

With four representatives among five successively majority species, *Firmicutes* phylum is largely dominant. Moreover, these four species belong to *Clostridia* class and three of them to *Clostridiales* order. The isolation of these strict anaerobic microorganisms from composting is in agreement with Martins et al. (2013) and emphasizes the fact that

anaerobic species play a key role in biomass degradation (Schloss et al., 2003) even in aerobic composting process.

Moreover, different parameters were calculated (Table 2) in order to provide more insights at the level of the ecological interpretation of the sequencing results. The number of observed OTUs is directly proportional to community richness. Starting from an initial value of 29, richness declines during the 43 first hours to attain a minimal value of 19. Afterwards, richness increases and reaches a final maximal value of 30 observed OTUs. By comparison with microbial communities typically found in compost, presenting high complexity with a large number of different taxa (Gladden et al., 2011), our consortium richness is very weak, suggesting that adaptation to selective conditions found in our experimental procedures induces a significant decrease of microbial diversity. The community isolated from our growth conditions is also characterized by its high level of unevenness expressed by Shannon's equitability index. Only few phylotypes among all observed OTUs are successively present in significant relative abundance. During the whole process, species distribution is relatively uneven with values between 0.26 and 0.58. The evolution of the environmental conditions induces a strong modification of population evenness. Indeed, its initial value is equal to 0.33, and the first 43 hours of growth are characterized by a slight increase of this parameter, the maximal value of the Shannon's equitability value being reached after 91 hours of growth. Then, evenness significantly decreases and reached its minimal value at the end of the process. From a theoretical point of view, evenness strongly influences the stability of the ecosystem, and is a very important parameter to be taken into account in order to microbial resource management (Wittebolle et al., 2009). High evenness provides higher probability that a species tolerant to perturbation is present in significant proportion, leading to an increase of the process robustness. However, despite the low evenness value

observed in our operating conditions, the consortium is able to display adaptation to the environmental fluctuations met during the cultivation tests (mainly due to pH variation). Deep modifications in community structure allow conserving cellulose degradation ability. These modifications, in terms of phylotypes' relative abundance, are quantified by means of Pearson product-moment correlation. The values (i.e., change in microbial composition) are calculated for each sampling time by comparison with the state observed at the inoculation (time = 0h), while rate of change expresses how rapidly the population evolves between two samples. Total change occurring between the beginning and the end of the process is equal to 96.0%. However, "change" and "rate of change" values show that modification does not occur uniformly during the process. During the first 43 hours, low change (3.4% after 43 hours) and weak rates of change are measured ($0.01\%.h^{-1}$ and $0.17\%.h^{-1}$). Fifty hours later, change is equal to 40.6% and a rate of change of $0.76\%.h^{-1}$ is calculated for the period between 43 hours and 91 hours of growth. Finally, period between 91 hours and 211 hours presents similar rate of change ($0.70\%.h^{-1}$).

Table 2 Ecological parameters values during anaerobic thermophilic (55°C) degradation of cellulose (filter paper 10 g.l^{-1}) by isolated consortium (10% v/v). Data for the observed OTUs, Shannon's equitability, Change (%) and Rate of change (%.h^{-1}).

Sampling (hours)	0	19	43	91	211
Observed OTUs	29	25	19	24	30
Shannon's equitability	0.33	0.39	0.38	0.58	0.26
Change (%)	0.0	0.2	3.4	40.6	96.0
Rate of change (%.h^{-1})		0.01	0.17	0.76	0.70

4 Conclusion

Based on microbial resource management, present work led to the isolation of a cellulolytic consortium able to boost anaerobic digestion. Important question behind this approach is the fate of exogenous microbial community during process. 16S rDNA sequencing pointed out weak richness of community. During cellulose anaerobic degradation, pH drop induced strong modification of microbial population. Despite weak richness and evenness, the consortium was able to adapt and to maintain cellulolytic potential. This important result point out the fact that simplified or synthetic microbial communities (De Roy et al., 2014) could be used in order to increase process robustness.

Acknowledgements

References

1. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. 1990. Basic local alignment search tool. *J Mol Biol*, **215**(3), 403-10.

2. Brosius, J., Dull, T.J., Sleeter, D.D., Noller, H.F. 1981. Gene organization and primary structure of a ribosomal RNA operon from Escherichia coli. *Journal of Molecular Biology*, **148**(2), 107-127.

3. Buswell, A.M., Mueller, H.F. 1952. Mechanism of Methane Fermentation. *Industrial & Engineering Chemistry*, **44**(3), 550-552.

4. Chao, A., Bunge, J. 2002. Estimating the number of species in a stochastic abundance model. *Biometrics*, **58**(3), 531-9.

5. Chao, A., Shen, T.-J. 2003. Nonparametric estimation of Shannon's index of diversity when there are unseen species in sample. *Environmental and Ecological Statistics*, **10**(4), 429-443.

6. Crespo, C., Pozzo, T., Karlsson, E.N., Alvarez, M.T., Mattiasson, B. 2012. Caloramator boliviensis sp. nov., a thermophilic, ethanol-producing bacterium isolated from a hot spring. *Int J Syst Evol Microbiol*, **62**(Pt 7), 1679-86.

7. De Roy, K., Marzorati, M., Van den Abbeele, P., Van de Wiele, T., Boon, N. 2014. Synthetic microbial ecosystems: an exciting tool to understand and apply microbial communities. *Environmental Microbiology*, **16**(6), 1472-1481.

8. Delcenserie, V., Taminiau, B., Delhalle, L., Nezer, C., Doyen, P., Crevecoeur, S., Roussey, D., Korsak, N., Daube, G. 2014. Microbiota characterization of a Belgian protected designation of origin cheese, Herve cheese, using metagenomic analysis. *J Dairy Sci*, **97**(10), 6046-56.

9. Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R. 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, **27**(16), 2194-2200.

10. Feng, Y., Yu, Y., Wang, X., Qu, Y., Li, D., He, W., Kim, B.H. 2011. Degradation of raw corn stover powder (RCSP) by an enriched microbial consortium and its community structure. *Bioresour Technol*, **102**(2), 742-7.

11. Gladden, J.M., Allgaier, M., Miller, C.S., Hazen, T.C., VanderGheynst, J.S., Hugenholtz, P., Simmons, B.A., Singer, S.W. 2011. Glycoside hydrolase activities of thermophilic bacterial consortia adapted to switchgrass. *Appl Environ Microbiol*, **77**(16), 5804-12.

12. Izquierdo, J.A., Sizova, M.V., Lynd, L.R. 2010. Diversity of bacteria and glycosyl hydrolase family 48 genes in cellulolytic consortia enriched from thermophilic biocompost. *Appl Environ Microbiol*, **76**(11), 3545-53.

13. Jabari, L., Gannoun, H., Cayol, J.-L., Hamdi, M., Fauque, G., Ollivier, B., Fardeau, M.-L. 2012. Characterization of Defluviitalea saccharophila gen. nov., sp. nov., a thermophilic bacterium isolated from an upflow anaerobic filter treating abattoir wastewaters, and proposal of Defluviitaleaceae fam. nov. *International Journal of Systematic and Evolutionary Microbiology*, **62**(Pt 3), 550-555.

14. Liu, Y., Yu, P., Song, X., Qu, Y. 2008. Hydrogen production from cellulose by co-culture of Clostridium thermocellum JN4 and Thermoanaerobacterium thermosaccharolyticum GD17. *International Journal of Hydrogen Energy*, **33**(12), 2927-2933.

15. Martins, L.F., Antunes, L.P., Pascon, R.C., de Oliveira, J.C.F., Digiampietri, L.A., Barbosa, D., Peixoto, B.M., Vallim, M.A., Viana-Niero, C., Ostroski, E.H., Telles, G.P., Dias, Z., da Cruz, J.B., Juliano, L., Verjovski-Almeida, S., da Silva, A.M., Setubal, J.C. 2013. Metagenomic Analysis of a Tropical Composting Operation at the São Paulo Zoo Park Reveals Diversity of Biomass Degradation Functions and Organisms. *PLoS ONE*, **8**(4), e61928.

16. 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**(6), 1571-81.

17. Monlau, F., Barakat, A., Steyer, J.P., Carrere, H. 2012. Comparison of seven types of thermo-chemical pretreatments on the structural features and anaerobic digestion of sunflower stalks. *Bioresource Technology*, **120**(0), 241-247.

18. Palmqvist, E., Hahn-Hägerdal, B. 2000. Fermentation of lignocellulosic hydrolysates. II: inhibitors and mechanisms of inhibition. *Bioresource Technology*, **74**(1), 25-33.

19. Pommier, S., Llamas, A.M., Lefebvre, X. 2010. Analysis of the outcome of shredding pretreatment on the anaerobic biodegradability of paper and cardboard materials. *Bioresource Technology*, **101**(2), 463-468.

20. Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., Glöckner, F.O. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research*, **35**(21), 7188-7196.

21. Ren, N., Cao, G., Wang, A., Lee, D.-J., Guo, W., Zhu, Y. 2008. Dark fermentation of xylose and glucose mix using isolated Thermoanaerobacterium thermosaccharolyticum W16. *International Journal of Hydrogen Energy*, **33**(21), 6124-6132.

22. Salehian, P., Karimi, K., Zilouei, H., Jeihanipour, A. 2013. Improvement of biogas production from pine wood by alkali pretreatment. *Fuel*, **106**(0), 484-489.

23. Schloss, P.D., Hay, A.G., Wilson, D.B., Walker, L.P. 2003. Tracking temporal changes of bacterial community fingerprints during the initial stages of composting. *FEMS Microbiol Ecol*, **46**(1), 1-9.

24. Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., Weber, C.F. 2009. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Applied and Environmental Microbiology*, **75**(23), 7537-7541.

25. Tachaapaikoon, C., Kosugi, A., Pason, P., Waeonukul, R., Ratanakhanokchai, K., Kyu, K.L., Arai, T., Murata, Y., Mori, Y. 2012. Isolation and characterization of a new cellulosome-producing Clostridium thermocellum strain. *Biodegradation*, **23**(1), 57-68.

26. Vavilin, V.A., Fernandez, B., Palatsi, J., Flotats, X. 2008. Hydrolysis kinetics in anaerobic degradation of particulate organic material: An overview. *Waste Management*, **28**(6), 939-951.

27. Verstraete, W., Wittebolle, L., Heylen, K., Vanparys, B., de Vos, P., van de Wiele, T., Boon, N. 2007. Microbial Resource Management: The Road To Go for Environmental Biotechnology. *Engineering in Life Sciences*, **7**(2), 117-126.

28. Viravaidya, K., Tangjerdjaras, W., Tachaapaikoon, C. 2011. Cellulosic bioethanol from clostridium thermocellum fermentation. in: *2011 International Conference on Chemistry and Chemical Process*, Vol. 10, IACSIT Press.

29. Wang, Y.-S., Byrd, C.S., Barlaz, M.A. 1994. Anaerobic biodegradability of cellulose and hemicellulose in excavated refuse samples using a biochemical methane potential assay. *J Ind Microbiol Biotechnol*, **13**(3), 147-153.

30. Wang, Y., Qian, P.-Y. 2009. Conservative Fragments in Bacterial 16S rRNA Genes and Primer Design for 16S Ribosomal DNA Amplicons in Metagenomic Studies. *PLoS ONE*, **4**(10), e7401.

31. Wittebolle, L., Marzorati, M., Clement, L., Balloi, A., Daffonchio, D., Heylen, K., De Vos, P., Verstraete, W., Boon, N. 2009. Initial community evenness favours functionality under selective stress. *Nature*, **458**(7238), 623-6.

32. Wongwilaiwalin, S., Rattanachomsri, U., Laothanachareon, T., Eurwilaichitr, L., Igarashi, Y., Champreda, V. 2010. Analysis of a thermophilic lignocellulose degrading microbial consortium and multi-species lignocellulolytic enzyme system. *Enzyme and Microbial Technology*, **47**(6), 283-290.

33. Yan, L., Gao, Y., Wang, Y., Liu, Q., Sun, Z., Fu, B., Wen, X., Cui, Z., Wang, W. 2012. Diversity of a mesophilic lignocellulolytic microbial consortium which is useful for enhancement of biogas production. *Bioresource Technology*, **111**(0), 49-54.

34. Zhang, Q., He, J., Tian, M., Mao, Z., Tang, L., Zhang, J., Zhang, H. 2011. Enhancement of methane production from cassava residues by biological pretreatment using a constructed microbial consortium. *Bioresource Technology*, **102**(19), 8899-8906.

35. Zuroff, T.R., Curtis, W.R. 2012. Developing symbiotic consortia for lignocellulosic biofuel production. *Appl Microbiol Biotechnol*, **93**(4), 1423-35.

CHAPTER II:

FLOW CYTOMETRY FINGERPRINTING – AN EFFICIENT LINK WITH MRM PARAMETERS

Abstract

Flow cytometry (FCM) is a high throughput single cell technology that is actually becoming widely used for studying phenotypic and genotypic diversity among microbial communities. This technology is considered in this work for the assessment of a bioaugmentation treatment in order to enhance cellulolytic potential of landfill leachate. The experimental results reveal the relevant increase of leachate cellulolytic potential due to bioaugmentation. Cytometric monitoring of microbial dynamics along these assays is then realized. The Flow FP package is used to establish microbial samples fingerprint from initial 2D cytometry histograms. This procedure allows highlighting microbial communities' variation along the assays. Cytometric and 16S rRNA gene sequencing fingerprinting methods are then compared. The two approaches give same evidence about microbial dynamics throughout digestion assay. There are however a lack of significant correlation between cytometric and amplicon sequencing fingerprint at genus or species level. Same phenotypical profiles of microbiota during assays matched to several 16S rRNA gene sequencing ones. Flow cytometry fingerprinting can thus be considered as a promising routine on-site method suitable for the detection of stability/variation/disturbance of complex microbial communities involved in bioprocesses.

Key-words: flow cytometry, leachate, landfill, cellulose, anaerobic digestion, bioaugmentation, amplicon sequencing.

80

1 Introduction

During the last decades, sanitary landfilling was the main technology used for the disposal of municipal solid waste (MSW) in Belgium. The sanitary landfills are designed for limiting the rainfall infiltration through the waste mass and collecting liquid and gaseous emissions from it (Benbelkacem et al., 2010). The low initial water content of MSW (20-25 %) and the low rate of rainfall infiltration within the landfill lead to relatively low biodegradation processes and an extended time for achieving long term stabilization of the MSW (Benbelkacem et al., 2010).

In order to accelerate the stabilization of the landfill, the idea of an 'anaerobic bioreactor landfill' was expressed 30 years ago (Hettiaratchi et al., 2015). This approach consists in the recirculation of leachate and the increase of MSW moisture for promoting the microbial activities and the biodegradation of organic materials. It has been shown that this approach improves methane production and enhance the MSW stabilization by comparison with "sanitary landfills" technology (Pohland and Kang, 1975). The anaerobic digestion assays performed by Gurijala and Sulfita (1993), for example, showed a faster methanization (3-4 fold), with MSW samples whose water content is close to 50 % than those whose moisture content were 20-30 %. At this level, an extensive review of scientific, technical and economic considerations related to "bioreactor landfills" are proposed by Reinhart et al. (2002), Sethi et al. (2013) and Hettiaratchi et al. (2015).

However, even in optimal conditions, hydrolysis of lignocellulosic substrate is the rate-limiting step during anaerobic digestion of wastes characterized by high lignocellulosic plant fiber content (Vavilin et al., 2008). Consequently, as lignocellulosic materials constitute

between 40 % and 70 % of MSW in developing countries (Li et al., 2009), a significant part of the landfill content is not accessible to biodegradation. According to Barlaz et al. (1989), between 23 and 29 % of lignocellulose is not degraded. This recalcitrance to degradation induces a progressive decrease of biogas production rate and an increase of landfill post-closure care period.

Bioaugmentation technology, in a microbial resource management view, by adding an exogenous cellulolytic microbial community isolated from compost, has been considered in this work. Indeed, addition of a microbial community with specific activity can potentially improve the performance of the in-situ biological processes, and facilitate the establishment of specific species in microbial communities (Bouchez et al., 2000). Frequently used to speed up the removal of undesirable molecules from environment, bioaugmentation induces also positive impact on anaerobic digestion of lignocellulosic substrate. Kinet et al. (2015) and Peng et al. (2014) highlight an increase of methane production from lignocellulosic substrates after the complementation of anaerobic digester sludge with cellulose degrading microbial agents. The bioaugmentation agents comprised either single cellulolytic species (Peng et al., 2014) or enriched cellulolytic community originally isolated from compost (Kinet et al., 2015). Furthermore, in their studies, Scherer and Neumann also highlighted an improvement of anaerobic digester performance through supplementation with compost (Neumann and Scherer, 2011; Scherer and Neumann, 2013). Additionally to degradation and fermentative activities improvement, these authors underlie that compost community addition enhances hydrogenotroph methanogen activity which in turn improves global performance of anaerobic digester.

A fast and efficient monitoring of the microbial communities is crucial for an efficient management of the bioprocesses. The impact of adding exogenous microbiota on global community and its persistence in the environment have to be assessed. Molecular tools, such as denaturing gradient gel electrophoresis (DGGE), are powerful for obtaining fingerprints of complex microbial communities (Marzorati et al., 2008). Moreover, next-generation sequencing methods, based on 16S rRNA gene sequencing, can provide information about the composition of complex microbial communities from various environments in a short time (Vanwonterghem et al., 2014). However, despite progresses in sequencing technologies have dramatically decreased the cost and increased the yield of sequence data generated (Vanwonterghem et al., 2014), 16S rRNA gene sequencing cannot be considered as a suitable solution for routine analysis of microbial communities. As a microbial community can rapidly adapt in response to environmental fluctuations, a reproducible sampling and fast analysis are required for managing community dynamics. Flow cytometry (FCM) is actually becoming widely used for studying phenotypic diversification among microbial isogenic communities (Delvigne et al., 2014; Delvigne and Goffin, 2014), and has been adapted here for the analysis of communities. Through the measurement of a set of optical parameters, FCM is able to characterize cellular populations at single-cell level in a very short time. Typically, intrinsic optical parameters give structural information about cells size and their internal organization whereas extrinsic parameters, resulting from a fluorescent labelling, provide supplementary physiological or functional cell properties. Particularly, cellular DNA content is a suitable target for analysis of complex communities. Sybr ® Green I and Syto9 ® dyes stain indifferently entire nucleic acid content whereas 4',6'-diamidino-2-phénylindole (DAPI) links specifically to adenine (A) and thymine (T) nucleic bases. Therefore, these dyes discriminate subcommunities according to their

respective DNA content (or A/T content). Moreover, several informatics tools (reviewed by Koch et al., 2014) have been developed for facilitating the interpretation of flow cytometric data. The establishment of flow cytometric fingerprints simplifies subsequent statistical analysis and particularly quantification of differences between samples. Koch and co-workers proposed a full flow cytometric workflow procedure (from sampling to data interpretation) for analyzing microbial intracommunity structure variation (Koch et al., 2013). This procedure, implemented with single dye staining (DAPI), has been used for the analysis of structural variations among a microbial community present in a biogas reactor (Koch et al., 2013b).

To our knowledge, only few publications deal with bioaugmentation for enhancing anaerobic digestion in landfill (Liu et al., 2013). The impact of the addition of a consortium, isolated from compostable matters, on leachate cellulolytic potential is assessed in this work. Furthermore, combination of 16S rRNA gene sequencing and flow cytometry approaches has been considered for monitoring microbial communities' dynamics throughout anaerobic digestion assays. More precisely, flow cytometry fingerprinting has been successfully used for assessing the stability of the complex microbial communities involved in anaerobic digestion.

2 Materials and methods

2.1 Biological matters

2.1.1 Compost

Approximatively 5 kg of composting matter are collected from a composting plant in Naninne, Belgium (50°25'37.7'' N, 4°54'27.7'' E), exclusively supplied with green lignocellulosic waste (30 000 ton treated per year). After grinding, rough ligneous fraction is separated from the fine fraction. Only this fine fraction is valorized through composting treatment. After moisture increasing, organic matters are aerated during four to six weeks using a blower to supply air from the bottom of the pile. Continuous monitoring of the oxygen concentration into pile ensures maintaining optimal conditions. Finally, transformed organic matters are matured during six to eight weeks. Matters are sampled at the end of forced aeration procedure before the maturation step. Temperature of ~70°C is recorded in pile during sampling.

2.1.2 Leachate

Leachate, used as inoculum for different tests, originates from landfill "CETeM" located in Mont-Saint-Guibert, Belgium (50°38'55.6'' N, 4°36'37.5'' E). Exploitation of this sanitary landfill began 55 years ago. 5.5 millions of tons of industrial and Municipal Solid Wastes (MSW) were buried there. Landfill gas has been valorized through electricity production since 1996. The leachates are collected at the bottom of the landfill and then treated in a Waste Water Treatment Plant (WWTP) for lowering Chemical Oxygen Demand (COD). Leachates from the WWTP are then re-injected into the landfill through an injection drain in order to increase the moisture of superficial MSW layers. Leachate sampling (10 L) is carried

out after the treatment step in WWTP. The dry matter content, total Kjedhal nitrogen concentration and COD of the leachate sample are respectively inferior to 1 g.L^{-1}, 0.5 g.L^{-1} and 20 g.L^{-1}. Volatile Fatty Acids (VFA) represent a minor part of COD content (<1 %).

2.1.3 Compost microbial consortium preparation

400 grams of fresh composting matters, packed in a nylon bag, are immerged in 10 L of pH 7 phosphate buffered saline (NaCl, 8 g.L^{-1}; KCl, 0.2 g.L^{-1}; Na$_2$HPO$_4$, 1.44 g.L^{-1}; KH$_2$PO$_4$ 0.24 g.L^{-1}) during 4 hours in a sterilized reactor (12 L) continuously stirred and flushed with oxygen-free nitrogen gas to install anaerobic atmosphere at room temperature. After four hours, the liquid phase containing biomass extracted from compost, is collected and concentrated ten times by centrifugation carried out at 4000.g.

2.2 Anaerobic digestion test

The cellulose digestion tests are carried out in triplicate for each type of inoculum in Biochemical Methane Potential (BMP) assay medium (Wang et al., 1994) supplemented with 10 g.L^{-1} of cellulosic material (Whatman filter paper) for 34 days at 55 °C under anaerobic condition in 125 ml sterile glass serum bottles without stirring.

Substrates are introduced into bottles containing 90 ml of BMP medium, and 10 ml of inocula. Inocula consist of either only leachate, extracted compost consortium or a mix 1:1 (vol:vol) of leachate and compost consortium. pH is adjusted with a 0.5 M KOH solution to achieve an initial pH of 7.3 in each sample, and a maximum variation during the culture period of pH ±1 is maintained. The sample bottles are capped tightly with rubber septa and sealed with aluminum seals. Nitrogen is passed into the bottles to flush out air and other gases before incubation.

The biogas productions are regularly collected using a syringe fitted with a needle. At these times, 5 ml of liquid phase are also collected for, metabolites (section 2.3), microbiota (section 2.4) and pH analysis.

Blank samples are realized to highlight the impact of organic matter introduced with the different inocula. Blank samples consist of 10 ml of inocula and 90 ml of BMP medium. No energetic substrate is added to the blank samples. Blank samples' biogas productions are subtracted from samples' production inoculated with the corresponding inoculum

At the end of the test, degradation ratios are calculated. Centrifugation (8000.g) allows the separation of supernatant and pellet which includes residual substrate and microbial biomass. Pellet is then suspended in 100 ml acetic acid/nitric acid reagent and heated at 100 °C for 30 min to remove the biological cells. Then, acetic acid/nitric acid treated suspension is filtered. The remaining cellulosic material is washed three times using 100 ml of distilled water each time. After washing and filtration, filtered solids are dried at 105 °C and determined gravimetrically.The weight loss of cellulosic materials is calculated by subtracting the weight of residual substrates from the total weight of cellulosic materials before degradation. The degradation ratio is calculated according to the equation:

$$Degradation\ ratio\ (\%) = \frac{M_t - M_r}{M_t} \times 100 \qquad (eq.1)$$

where M_t is the total weight of the cellulosic materials before the degradation and M_r is the weight of residual substrates after the degradation.

2.3 Metabolites analysis

Volatile fatty acids, ethanol and glucose concentrations in liquid phase are determined using an Agilent 1110 series high performance liquid chromatograph (HPLC; Agilent Technologies)

equipped with a Supelcogel C-610H column (Sigma-Aldrich, St Louis, MO, USA) preceded by a Supelguard H precolumn (Sigma-Aldrich). The columns are kept at a temperature of 40 °C, and the isocratic mobile phase is 0.1 % H3PO4 (in ultrapure, ''milliQ'', water), at a flow rate of 0.5 ml.min^{-1}. A differential refraction index detector, kept at 35 °C, is used. This analysis takes 35 min at a maximum pressure of 60 bars.

2.4   Microbial communities analysis

2.4.1   Molecular analysis of microbial communities

Total DNA extraction

Total DNA is isolated from each primary suspension using the DNeasy Blood & Tissue DNA extraction kit (Qiagen Benelux B.V., Venlo, the Netherlands), following the manufacturer's extraction for gram-positive bacteria. Briefly, the bacteria present in the suspension are lysed for 1 hour at 37 °C in a lysis buffer (Tris-HCL 20 mM, sodium EDTA 2 mM, triton x-100 1.2 % and lysozyme 20 mg.ml^{-1}). The suspension is then treated with proteinase K for 1 hour at 56 °C. The resulting suspension is mixed with ethanol and loaded on a Dneasy Mini-spin column. The DNA is then washed 2 times with QIAgen buffers and eluated into DNase/RNase-free water and its concentration and purity are evaluated by optical density using the NanoDrop ND-1000 spectrophotometer (Isogen Life Science NV, Sint-Pieters-Leeuw, Belgium). The DNA samples are stored at −20 °C until use in 16S rRNA gene amplicon pyrosequencing analysis.

16S rRNA Gene Library Construction and Pyrosequencing

16S rRNA Gene PCR libraries are generated for the samples using 16S rRNA gene universal primers with Illumina overhand adapters targeting V1-V3 hypervariable regions, forward (5'-

TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3') and reverse (5'-GTVTVGTGGGCTCGGAGATGTGTATAAGAGACAG-3'). Each PCR product is purified with the Agencourt AMPure XP beads kit (Beckman Coulter, Pasadena, USA) and submitted to a second PCR round for indexing, using the Nextera XT index primers 1 and 2. After purification, PCR products are quantified using the Quant-IT PicoGreen (ThermoFisher Scientific, Waltham, USA) and diluted to 10 ng.µL^{-1}. A final quantification, by qPCR, of each sample in the library is performed using the KAPA SYBR® FAST qPCR Kit (KapaBiosystems, Wilmington, USA) before normalization, pooling and sequencing on a MiSeq sequencer using v3 reagents (ILLUMINA, USA).

16S rRNA gene Data Processing

Sequence reads processing are used as previously described (Rodriguez et al., 2015) using respectively MOTHUR software package v1.35 (Schloss et al., 2009), Pyronoise algorithm and UCHIME algorithm (Edgar et al., 2011) for alignment and clustering, denoising and chimera detection. 16S Reference alignment and taxonomical assignation are based upon the SILVA database (v1.15) of full-length 16S rRNA gene sequences. The final reads are clustered into operational taxonomic units (**OTUs**) with the nearest neighbor algorithm using MOTHUR with a 0.03 distance unit cutoff. A taxonomic identity is attributed to each OTU by comparison with the SILVA database (80% homogeneity cutoff) (Delcenserie et al., 2014). When taxonomic identification fell below the 80 % threshold, the taxonomic level was labelled with the first defined level from higher level followed by the term "_unclassified". All unique sequences for each OTU were further compared with the SILVA data set version v1.15 using the BLASTN algorithm (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_

LOC=blasthome (Altschul et al., 1990), as MOTHUR is not suitable to taxonomic assignment beyond the genus level. For each OTU, a consensus detailed taxonomic identification is given based upon the identity (less than 1% of mismatches with the aligned sequence) and the metadata associates with the most frequent hits (validated bacterial species or not).

All the biosample raw reads have been deposited at the National Center for Biotechnology Information (NCBI) and are available under de Bioproject ID PRJNA315623.

Ecological parameters calculation

Relative abundance of every OTUs (at genus level) detected in different samples have been determined for each samples. This results for each sample in a vector of values. Based on these vectors of values, inter-samples Pearson's product moment correlation coefficients have been calculated.

2.4.2   Flow cytometry fingerprinting

Flow cytometry analysis

FCM analyses are realized directly after sampling to avoid community evolution. They are carried out on BD Accuri C6 flow cytometer equipped with 20 mW Solid State Blue Laser (488 nm) and 14.7 mW Diode Red Laser (640 nm). Prior to analysis, the samples are filtered with 30 μm cut-off filter paper in order to eliminate insoluble particles. After filtration, the microbial biomass is suspended in pH 7 phosphate buffered saline (NaCl, 8 g.L^{-1}; KCl, 0.2 g.L^{-1}; Na_2HPO_4, 1.44 g.L^{-1}; KH_2PO_4 0.24 g.L^{-1}) and the biomass concentration is adjusted to reach approximatively 10^6 cells per ml. All-cell labeling approach based on DNA labelling is then realized by using SYBR® Green I. SYBR® Green I stock solution (10,000 times concentrate, Life Technologies) is diluted 100 times in dimethyl sulfoxide (DMSO) and filtered on 0.22 μm. 1

ml of cell suspension is supplemented with 10 µl of stain solution and incubated 30 minutes at room temperature. Forward light Scatter (FSC), Side light Scatter (SSC) are collected with a 488/10 bandpass filter and green fluorescence (FL1 channel) is collected with photomultiplier tubes using 533/30 bandpass filter. For each sample run, data for 40 000 events are collected with a flow rate of 14 µl.min^{-1}.

Data analysis

The raw data are extracted from the CFlow software (Accuri, BD Bioscience) as .fcs files and loaded into R software. Data analysis pipeline (supplementary file 1) is adapted from the version developed by De Roy et al. (2012) and consists in three principal steps. (1) Creating a n-dimensional quantitative fingerprint of each sample from their respective bivariate (SSC – FL1) FCM distribution using the recursive probability binning (PB) algorithm for flow cytometry data, implemented in the Bioconductor package FlowFP (Rogers and Holyst, 2009). In a first time, a model, composed of hyper-rectangular regions (bins) of varying size and shape, is established. From superimposed FCM distribution (i.e. data from all samples are pooled together) and thanks to Probability Binning (PB) algorithm (Rogers et al., 2008), bivariate data space is divided in hyper-rectangular regions in such way that each contains similar number of events (one event corresponding to one cell). First step of the algorithm consists in the division of space into two bins containing similar number of events. Afterwards, each of these bins is again divided into two bins with equal number of events, and so forth. Therefore region of bivariate FCM space displaying high density of events are characterized by bins of small area whereas larger bins characterized regions of weak density. Moreover, the final number of bins (n) is arbitrary set and correspond to 2^i with i the number of recursive subdivisions. In our case, 6 recursive subdivisions are applied and therefore model is composed of 64 bins. The obtained model is then applied to each sample

and number of cells per bin is determined, creating feature vector of counts (n-dimensional) for each sample. The latter will be used as a fingerprint of the microbial community at a given time and under given conditions. (2) Principal component analysis of the different fingerprints. The 64 bins are considered as the variables described for each sample by the number of cells per bin determined previously. The dimension reduction is established by selecting number of principal components so that 96 % of the initial variance is explained. (3) Calculation of Pearson's product moment correlation coefficients between samples according to their coordinates in principal components referential.

3 Results and Discussion

3.1 Bioaugmentation of leachate leads to the enhancement of cellulolytic potential of microbial communities

Bioaugmentation have been applied to leachate in order to enhance its cellulolytic potential. The addition of a microbial consortium isolated from composting plant is considered for this purpose. Assays of anaerobic digestion of cellulose have been carried out in order to assess the efficiency of this treatment. Three kinds of inocula have been used: (i) leachate from landfill, (ii) consortium isolated from compost and (iii) bioaugmented leachate obtained by the addition of the compost consortium to leachate with a volumetric ratio 1:1.

The cellulose hydrolysis rates, obtained after 34 days of digestion process, are presented at Figure 1.A. Microbial community from bioaugmented leachate is able to hydrolyze the whole cellulose content after 34 days (99 ± 0%), while respective hydrolysis rates observed for compost consortium alone and leachate alone are 65 ± 16 % and 58 ± 36 %. These results

clearly point out the positive impact of the bioaugmentation treatment on the leachate's hydrolytic potential.

For improving the comprehension of this effect, different parameters have been recorded during cellulose digestion processes. Concentration of biodegradation products and biogas productions throughout the assays, presented at Figure 1.B and 1.D, have been used as proxy for estimating the kinetics of microbial activities. During the first days, the highest
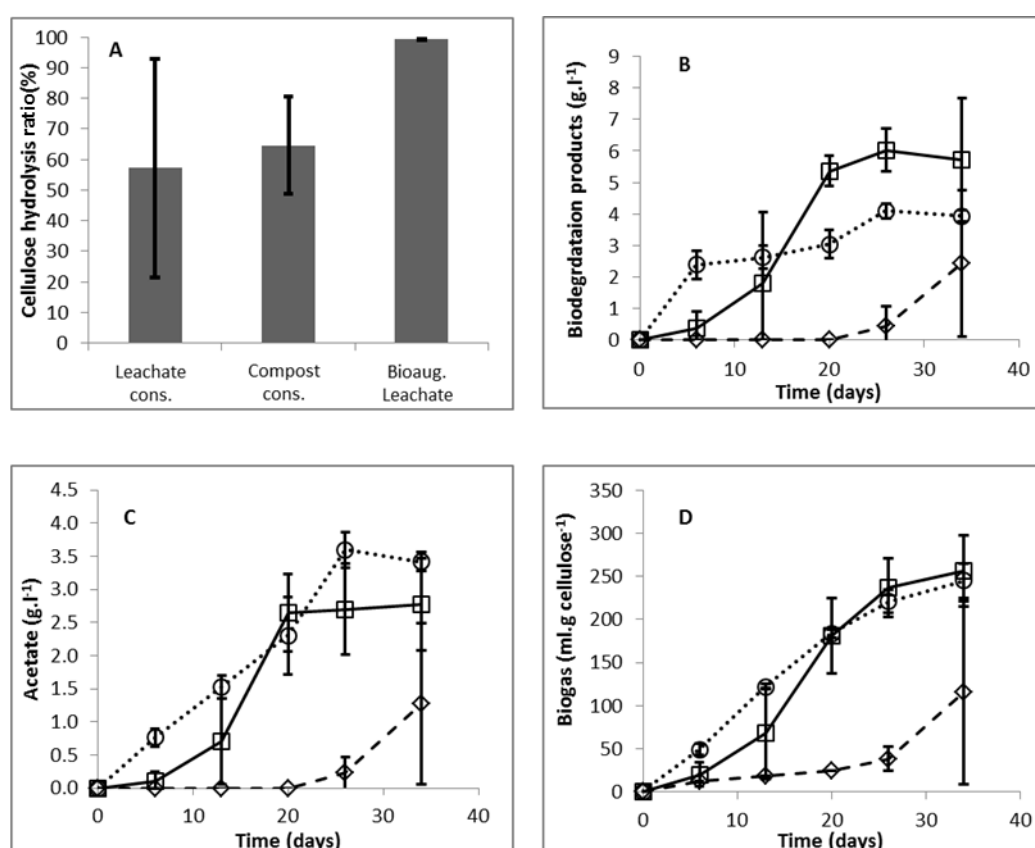

Figure 1 Experimental results of cellulose anaerobic digestion assay in triplicate using 3 different inocula: leachate (◊), compost consortium (○) and bioaugmented leachate (□). The curves represent: A-final cellulose hydrolysis rate; B-Evolution of biodegradation products concentration throughout the process ; C- Evolution of acetate concentration throughout the process and D-Cumulative biogas production throughout the process.

productions are observed in digestion flasks containing compost consortium. After 6 days, compost consortium already produces 2.39 ± 0.46 g.L^{-1} of biodegradation products and 49 ± 7 ml of biogas whereas weak productions are recorded for bioaugmented leachate and leachate. Significant productions are measured after respectively 26 days and 13 days for

leachate and bioaugmented leachate microbial communities. Compost consortium is directly able to efficiently metabolize cellulose whereas bioaugmented leachate and leachate need adaptation period. Adaptation period observed in the case of bioaugmented leachate is probably due to a dilution of the cellulolytic bacteria. Afterwards, synergistic action of microbial communities coming from both environments induces higher microbial activities. After 20 days, cumulative biogas production by bioaugmented leachate reaches similar value than the one measured for compost consortium and the concentration of biodegradation products exceeds concentration obtained with compost consortium. Finally, between the 20th day and the 34th day, biogas production by bioaugmented leachate overcomes biogas production by compost consortium. Concerning concentrations of biodegradation products, highest values are still measured in bioaugmented leachate samples after 34 days. Total concentrations of biodegradation products in liquid phase shower the same trend (i.e. a slight increase followed by stabilization) for bioaugmented leachate and compost consortium microbial communities. On the opposite, specific acetate concentrations (Figure 1.C) evolve differently during this last period for these two microbial communities. Acetate concentrations in compost consortium samples still increase and reach values superior to the typical inhibitory concentration (3 g.L^{-1}) while stability is observed for bioaugmented leachate. This rise above the critical threshold, probably due to the lack of efficient acetate consumers in compost microbial community, could explain the lower final cellulolytic hydrolysis rate (Figure 1.A) observed with compost consortium inoculum. In the case of bioaugmented leachate, this deficiency seems to be offset by bacteria originally present in leachate microbiota. It is also important to notice that biogas and other metabolites productions are observed only for one replicate of digestion assays led with not modified

leachate microbial community. This explains the high variability of the values observed after 34 days (Figure 1).

All these results point out that bioaugmentation treatment accelerates cellulose digestion process through the fast establishment of a cellulolytic community probably originating from compost consortium. Moreover, synergistic action between endogenous and exogenous microbial communities allows avoiding inhibitory conditions and leads to a complete assimilation of cellulose. On the opposite, a longer time is required for the establishment of an active cellulolytic community in untreated leachate; and due to a lack of metabolic activities preventing the accumulation of inhibitory compounds such acetate, consortium isolated from composting plant is not able to sustain complete cellulose hydrolysis.

3.2 Microbial community analysis

Flow cytometry and 16S rRNA gene sequencing have been considered for the analysis of the structure and dynamics of microbial communities involved in this work. For each type of inoculum, both methods are carried out for the replicate expressing maximum final cellulose degradation ratio.

FCM fingerprinting reproducibility has been confirmed by comparing fingerprint obtained with the replicate displaying the highest cellulose degradation ratio and the average fingerprint resulting from the analysis of the communities from three replicates. More precisely, analysis of compost communities have been performed in triplicates and has been further used in this validation approach.

The evolution of biogas production and degradation products concentrations obtained for these replicates during cellulose digestion tests can be found at figure 2.
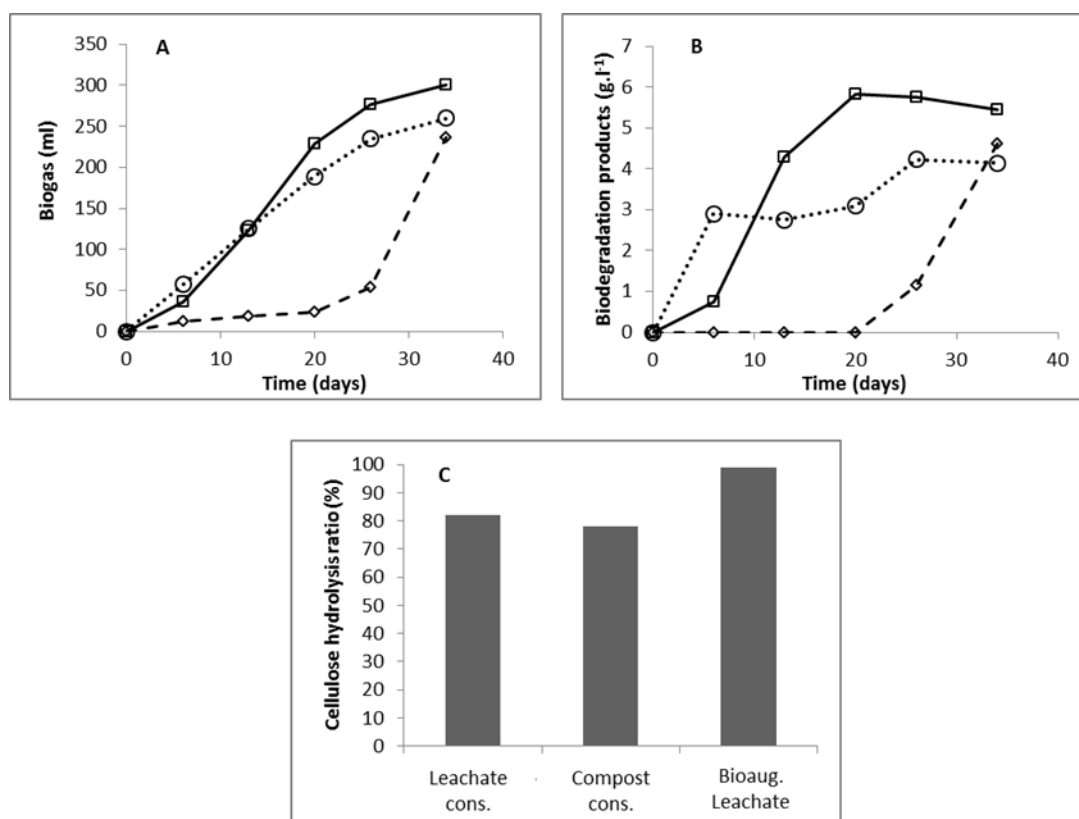
Figure 2 experimental results of cellulose anaerobic digestion assay using 3 different inocula: leachate (◊), compost consortium (○) and bioaugmented leachate (□). The curves represent: A-final cellulose hydrolysis rate; B- Evolution of biodegradation products concentration throughout the process ; C- Evolution of acetate concentration throughout the process and D-Cumulative biogas production throughout the process. The values presented here express, for each inoculum, the results obtained for the replicate displaying the maximal final cellulose degradation ratio

3.2.1 Bacterial diversity analysis based on 16S rRNA gene sequencing

The main bacterial genera (relative abundance superior to 5% in at least one sample) met in the different samples for different times during the anaerobic degradation process are shown at figure 3. Moreover, initial leachate and compost consortium microbial community composition are presented. Initially, these two communities are composed by respectively 61 and 179 OTUs (data not shown). Afterwards, important decrease of richness occurs. At the end of the digestion process, 19 OTUs, 11 OTUs and 13 OTUs are respectively observed in compost consortium, leachate and bioaugmented leachate microbial communities (data not shown).
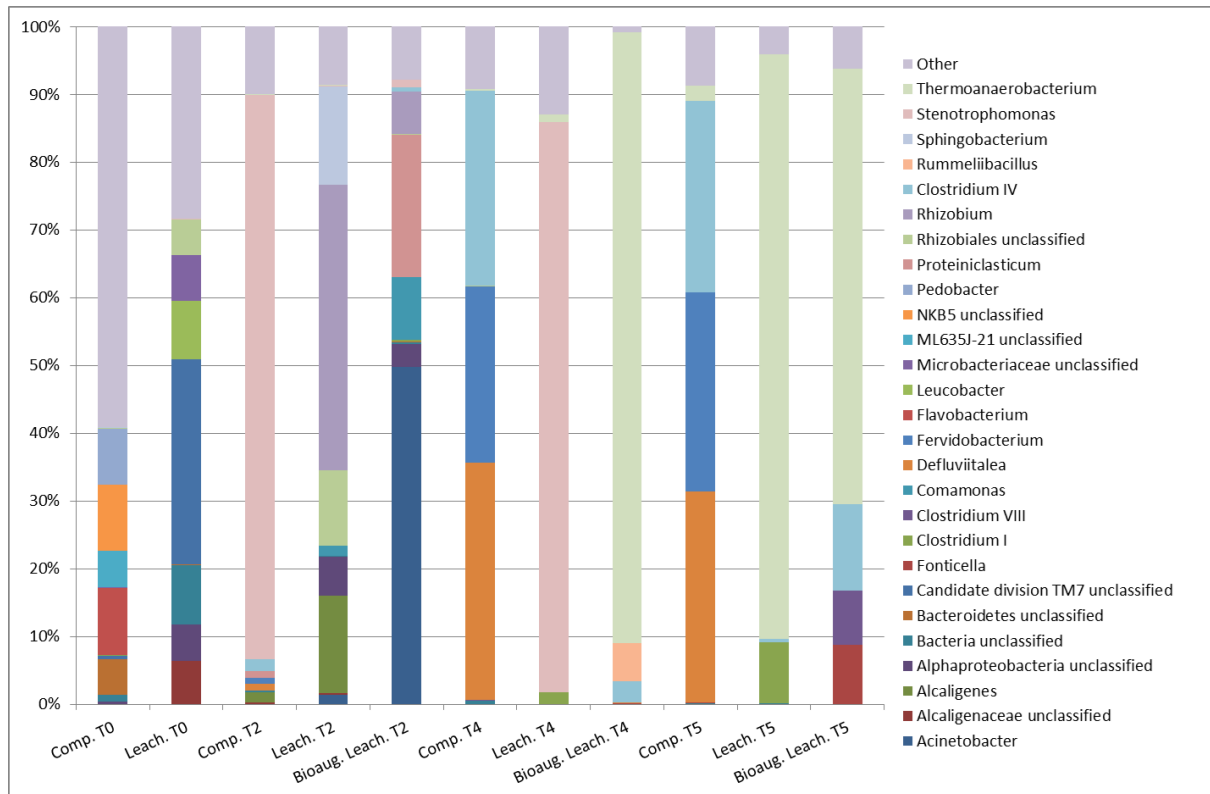
Figure 3 16S rRNA gene sequencing fingerprint of samples from 3 kinds of inocula: (1) compost consortium (Comp.), (2) leachate (Leach.), (3) bioaugmented leachate (Bioaug. Leach.) during cellulose (filter paper, 10g.l^{-1}) anaerobic and thermophilic (55°C) digestion. Samples are taken at 4 sampling times : T0 (0 day), T2 (13 days), T4 (26 days) and T5 (34 days). As bioaugmented leachate consist in a mix 1:1 (v:v) of compost consortium and leachate, sequencing has not been carried out on Bioaug. Leach. sample before the beginning of the anaerobic digestion (T0).

After 13 days of process, compost consortium is mainly composed of bacteria affiliated to the *Stenotrophomonas* genus (83%). Interestingly, bacteria belonging to this genus and more precisely to *Stenotrophomonas maltophila* species have already been reported to be able to degrade cellulose (Huang et al., 2012). This is the only OTU present at a relative abundance superior to 5%. However, two other potential cellulolytic OTUs are detected in weak proportion. The first one corresponds to bacteria which can be assigned to the *Defluviitalea* genus and closely relative to a bacterium already found in degrading cellulose community (EU250956). The second one is affiliated to *Clostridium* genus. Furthermore, the majority of the sequences grouped in this OTU can be assigned to *Clostridium cellulosi* species. This

species is recognized to be able to hydrolyze cellulose and to assimilate the released cellobiose in anaerobic and thermophilic conditions (He et al., 1991). These two genera present respective relative abundances of 1% and 2%. None of 5 OTUs, displaying a relative abundance superior to 5% in crude compost consortium, is still detected. Microbial community rapidly adapts to specific growth conditions (i.e. cellulose as sole carbon source). For their part, leachate and bioaugmented leachate display several OTU with relative abundance superior to 5% after the same period. Bacteria belonging to *Rhizobium* genus are present in proportion superior to 5% in both consortia. Relative abundance of 42% and 6%, are respectively determined in leachate and bioaugmented leachate. Bacteria from *Alcaligenes* genus are also present in significant proportion (14%) in leachate consortium. Uncultured bacterium (FN436139) affiliated to this genus has been previously detected in thermophilic biogas reactor fed with renewable biomass. Moreover, presence of bacteria affiliated to this genus has been highlighted in an efficient lignocellulose-degrading microbial consortium (Hui et al., 2013). Others main OTUs from leachate microbiota are assigned to *Sphingobacterium* genus, *Rhizobiales* order and *Alphaproteobacteria* class and present respective relative abundance of 14, 11 and 6%. Therefore some OTUs (i.e. Alphaproteobacteria and Rhizobiales) displaying non negligible relative abundance in crude leachate consortium are still present after 13 days of process. Contrary to compost consortium community, leachate microbiota seems slower to adapt to growth conditions. This observation is in agreement with results shown at figure 2 which highlight inefficient cellulose digestion by this consortium.

Bioaugmented leachate community is dominated (50%) by bacteria belonging to *Acinetobacter* genus. The second and third most abundant OTUs are respectively affiliated to *Proteiniclasticum* and *Comamonas* genera and display respective relative abundance of 21%

and 9%. Finally, it can be noticed that *Clostridium cellulosi* species display a relative abundance of 1% in bioaugmented leachate after 13 days of process. So despite efficient cellulose digestion by bioaugmented leachate community (figure 2), only few 16S rRNA gene sequences are assigned to strains known for their cellulolytic potential. It can be explained either by the lack of knowledge about metabolic characteristics of uncultured environmental bacteria either by amplification of inactive strains' sequences. Indeed, this technique does not allow making the distinction between died, dormant and active cells. Amplification of sequences from inactive cells can hide information related to the active bacteria.

On the opposite, after 26 days, as well in compost consortium as in bioaugmented leachate, a majority of sequences are assigned to genera, or even species, known for their cellulolytic potential. Bacteria belonging to *Clostridium cellulosi* and *Defluviitalea* genus, previously detected in small amount, display respective relative abundance of 29% and 35% in compost consortium. 16S rRNA gene sequencing also reveals that 26% of sequences belong to bacteria which can be affiliated to *Fervidobacterium* genus. In a previous study (Kinet et al., 2015), bacteria belonging to *Defluviitalea* and *Fervidobacterium* genera has already been isolated from compost and shown to be involved in anaerobic and thermophilic cellulose degradation. Concerning bioaugmented leachate community, 90% of analyzed sequences are assigned to previously not detected genus *Themoanaerobacterium*. 84% of these sequences can be even more affiliated to *Themoanaerobacterium thermosaccharolyticum* species which is known for its ability to ferment cellobiose. Moreover, O-Thong et al. (2008) and Ren et al. (2008) showed that some specific strains belonging to this species are able to metabolize cellulose. Contrary to what is happened for compost consortium, *Clostridium* genus still displays low relative abundance (3%). Finally, bacteria assigned to *Rummeliibacillus* genus represent 6% of total bioaugmented leachate community.

99

The efficient cellulose digestion observed for leachate microbial community between 20th and 34th days is mirrored by emergence and succession of new OTU in this community. After 26 days, microbiota is largely dominated by *Stenotrophomonas* genus (84%), which is the only OTU displaying relative abundance superior to 5%. On the opposite, after 34 days, Stenotrophomonas genus completely disappears from leachate microbiota whereas Thermoanaerobacterium genus becomes largely dominant (86%) and bacteria affiliated to *Clostridium* sensu stricto (cluster I) genus appeared in non-negligible proportion (9%). Notice that contrary to what is observed in bioaugmented leachate community, any sequence assigned to *Themoanaerobacterium* OTU can be affiliated to *Themoanaerobacterium thermosaccharolyticum* species.

Despite decrease between 26th and 34th days, *Themoanaerobacterium* genus is still dominant in final bioaugmented leachate microbiota with 64% of analyzed sequences assigned to this genus. Moreover, 59% of sequences are still affiliated to *Themoanaerobacterium thermosaccharolyticum*. Contrary, relative abundance of *Clostridium cellulosi* increases from 3% to 13%. Two other OTUs display significant relative abundance of respectively 9% and 8%. First one is affiliated with *Fonticella* genus. Bacteria belonging to this genus have already been identified in compost (uncultured bacterium (FN667330)) (Partanen et al., 2010). Second one is affiliated with *Clostridrium* sensu stricto (cluster VII) genus. Finally, only weak modifications affect compost consortium microbiota. Three main strains present after 26 days display pretty same relative abundance, 31%, 29% and 28% respectively for bacterium belonging to Defluviitalea genus, *Fervidobacterium* genus and *Clostridium cellulosi* species.

Except, one classified among *Fervidobacterium* genus, all types of strain present in proportion superior to 5% in different final communities are part of *Clostridia* class. This highlights the importance of this class of bacteria during cellulose anaerobic digestion process in thermophilic conditions.

Taken altogether, the results coming from sequencing analyses show the importance of the succession of microbial communities in the hydrolysis of cellulose and the subsequent metabolic activities leading to the generation of biogas. Therefore, a special attention must be paid to the dynamics of the communities in order to ensure the efficiency of the processes. As amplicon sequencing cannot be considered for routine on-site analysis, complementary cheaper and easier technique giving access to the dynamics of microbial communities is required. Flow cytometry fingerprinting has been considered for this purpose. More specifically, FC fingerprinting has been used in order to assess the stability of the communities and to assess resilience of consortium isolated from composting plant when inoculated in leachate.

3.2.2 Flow cytometry community fingerprinting for routine population monitoring

Flow cytometric analyses have been performed for microbiological samples whose 16S rRNA gene sequencing has already been performed. Note that, additionally to samples taken after 13 (T2), 26 (T4) and 34 days (T5), consortia after 20 days (T3) of process are also analyzed with flow cytometry. Basic flow cytometry dot plots (SSC-FL1) are shown at figure 4. As internal organization is more discriminating than cellular size, SSC parameter has been preferred to FSC parameter for the characterization of communities. It can be seen that different population structures can be considered. These differences will be parameterized

on the basis of FCM fingerprinting. Further principal component analysis has been applied to the different fingerprints. Distribution of the different samples according to the two main principal components is represented at figure 5.

Table 1 and Table 2 display correlation coefficients between samples according to flow cytometry and 16S rRNA gene sequencing fingerprinting. Amplicon sequencing based coefficients correlation are considered as reference for validating FCM based observation.

The evolution of leachates' cytometry fingerprints is different from two others (figure 5). After 20 days (T3), leachates' cytometry fingerprints are still displaying a strong evolution. On the opposite, fingerprints of compost and bioaugmented leachate communities stabilize. This is particularly relevant for the compost consortium. Compost and bioaugmented leachate consortia display respective correlation coefficients of 0.89 and 0.74 between successive 26th day (T4) and 34th day (T5) samples while coefficient of 0.16 is calculated for the leachate community. These results are in accordance with the dynamics of anaerobic digestion displayed in Figures 1B, 1C and 1D. The fastest stabilization of the microbial community is observed for the compost inoculum while the slowest one is observed for leachate inoculum. Interestingly, bioaugmentation treatment applied to leachate accelerates the establishment of a stable community adapted to specific growth conditions (i.e. community able to grow in anaerobically and with cellulose as only carbon source). Indeed, stabilization dynamics of microbial community is significantly accelerated for bioaugmented leachate compare to leachate inoculum.
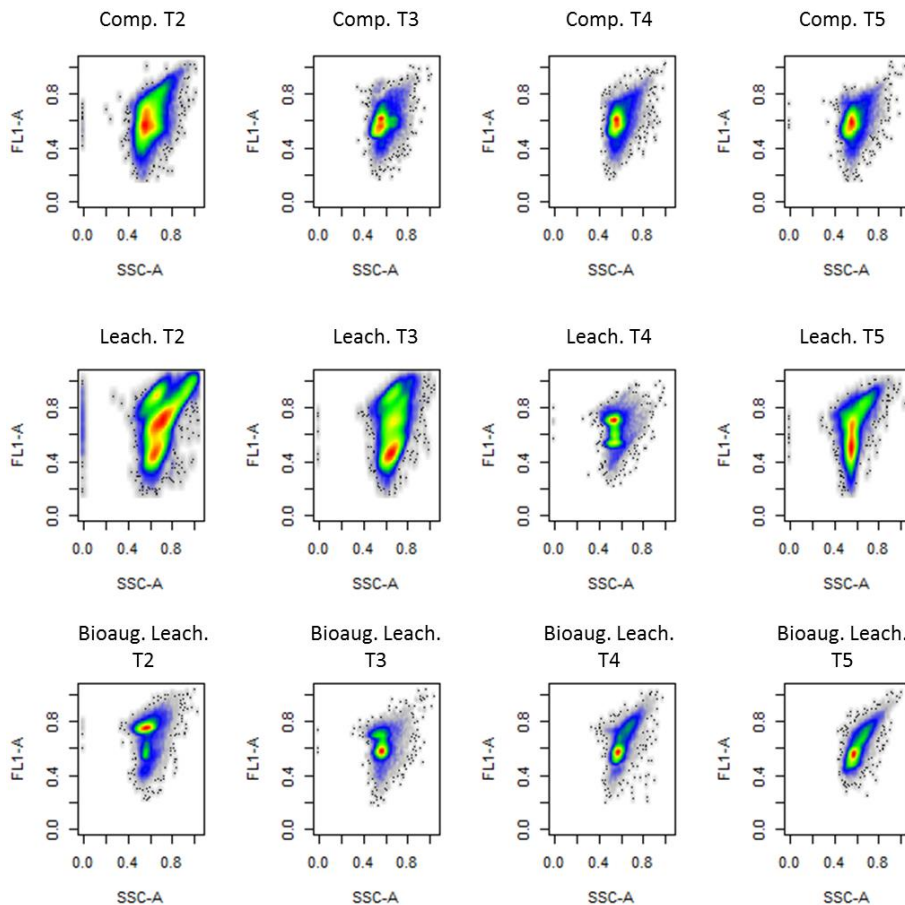
Figure 4 Cytogram (SSC-A/FL1-1) from 3 communities: compost consortium (comp.), leachate consortium (Leach.) and bioaugmented leachate (Bioaug. Leach.) during cellulose (filter paper, 10g.l⁻¹) anaerobic and thermophilic (55°C) digestion. Samples are taken at 4 sampling times: T2 (13 days), T3 (20 days), T4 (26 days) and T5 (34 days).

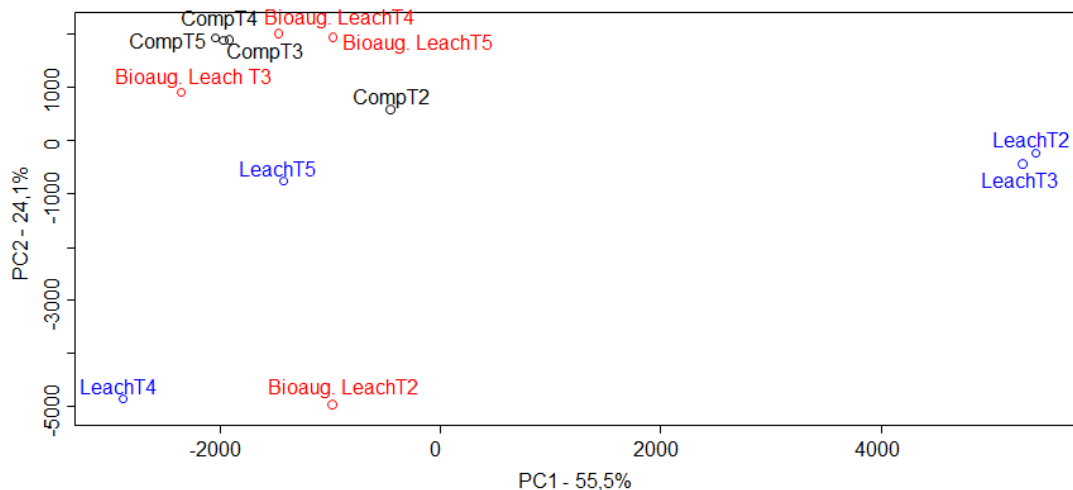**Figure 5 Evolution of microbiotas (Comp (compost consortium), Leach (leachate consortium), and Bioaug Leach (bioaugmented leachate)) flow cytometric fingerprints in 2 dimensions space (two first principal components). Microbiotas are sampled at 4 times during cellulose (filter paper 10 g.l⁻¹) anaerobic and thermophilic (55°C) digestion: T2 (13 days), T3 (20 days), T4 (26 days) and T5 (34 days).**

Table 1 assessment of microbiota stabilization for the 3 kinds of inocula: Compost consortium (Comp.), leachate consortium (Leach.) and bioaugmented leachate (Bioaug. Leach). Correlation between communities at successive sampling points (13 days (T2) – 26 days (T4) and 26 days – 34 days (T5)) are evaluated through Pearson's product moment correlation coefficients. These correlation coefficients are calculated based on 16S rRNA gene sequencing data and FCM fingerprinting data obtained for microbial communities monitoring.

Microbiota pairs	Amplicon sequencing	Flow cytometry
	Correlation coefficients	Correlation coefficients
Comp. 13 days – Comp. 26 days	-0.03	0.16
Comp. 26 days – Comp. 34 days	0.99	0.89
Leach. 13 days – Leach. 26 days	-0.06	-0.39
Leach. 26 days – Leach. 34 days	-0.02	0.16
Bioaug. Leach. 13 days – Bioaug. Leach. 26 days	-0.07	-0.03
Bioaug. Leach. 26 days – Bioaug. Leach. 34 days	0.96	0.74

Table 2 Comparison of microbiota from the 3 kinds of inocula (Leachate consortium (Leach), Compost consortium (Comp) and Bioaugmented leachate (Bioaug. Leach)) for 3 sampling times (T2-13 days, T4-26 days and T5-34 days) using Pearson's product moment correlation coefficient calculated according to 16S rRNA gene sequencing data and FCM fingerprinting data.

Time	Microbiota pairs	Amplicon sequencing	Flow cytometry
		Correlation coefficients	Correlation coefficients
13 days	Leach. – Comp.	-0.13	-0.27
	Leach. – Bioaug. Leach.	-0.06	-0.08
	Comp. – Bioaug. Leach.	-0.09	0.03
26 days	Leach. – Comp.	-0.11	-0.15
	Leach – Bioaug. Leach.	-0.06	-0.17
	Comp. – Bioaug. Leach.	-0.11	0.50
34 days	Leach. – Comp.	-0.08	-0.09
	Leach. – Bioaug. Leach.	0.96	0.05
	Comp. – Bioaug. Leach.	-0.001	0.39

These FCM based trends about stabilizations dynamics are in agreement with sequencing results described in previous section. Correlation coefficients, calculated based on successive sequencing fingerprints display similar trends than FCM based correlation coefficients.

Correlation coefficients obtained after the 26th day and the 34th day, are equal to 0.99 and 0.96 for respectively bioaugmented leachate and compost communities while value of -0.02 is observed for leachate consortium. Comparatively, correlation coefficients of -0.03, -0.07 and -0.06 are calculated between T2 and T4 for respectively compost, bioaugmented leachate and leachate consortia. This good agreement between cytometry fingerprints and sequencing data to assess microbial community stabilization is encouraging for the application of FCM in microbial resource management perspective and more precisely for the monitoring of the communities stability.

Otherwise, according to the distribution of FCM fingerprints in principal components space (figure 5), bioaugmented leachate microbiota is more similar to compost community than leachate consortium from the 20th day (T3) of the process. Based on FCM data, correlation coefficients of 0.50 and 0.39 are calculated between bioaugmented leachate and compost communities respectively after 26 and 34 days whereas leachate-bioaugmented leachate coefficient correlations exhibit values of -0.17 and 0.05 (Table 2). These results would reveal the potential impact of bioaugmentation treatment on the phenotypic profile of microbial communities. The growth of strains provided by cellulolytic consortium would lead to a reconciliation of respective cytometric fingerprints of compost consortium and bioaugmented leachate. However, these trends are contradicted by sequencing results. Amplicon sequencing based correlation coefficients of -0.11 and -0.001 between compost and bioaugmented leachate microbiota are calculated after respectively 26 and 34 days of process. In this case, flow cytometry fingerprinting cannot adequately discriminate microbial samples composed of different bacterial genera. Similar cytometry fingerprints are obtained for different amplicon sequencing profiles. This can be explained by the fact that different

types of bacteria can display similar flow cytometric characteristics, as shown in works of Koch et al. (2013) and Müller et al. (2012). This is actually one of the main drawbacks of this cytometric fingerprinting method.

Flow cytometry fingerprinting is a promising approach for dense and low cost on-site monitoring of microbial communities involved in bioprocesses such as anaerobic digestion. However, this technique can only be considered as a first line method giving indication about the stability/disturbance of an established microbial community along the bioprocess, and cannot be dissociated from 16S rRNA gene sequencing when more precise community characterization is required. Indeed, staining technique implemented in our flow cytometric approach is not specific enough for fine discrimination of the different genera/species involved in the community. Further improvement of the staining technique must be considered. FISH combined with flow cytometry can also be an alternative for the discrimination of the main constituent of the microbiota.

Interestingly, as already successfully implemented for characterizing heterogeneities in isogenic population bioreactor (Brognaux et al., 2013), flow cytometry could be considered for automated on-line monitoring of complex population from anaerobic digestion.

4 Conclusion

Bioaugmentation of leachate microbiota with consortium isolated from composting plant induces a significant improvement of leachate cellulolytic potential (average improvement of 41 %). Combination of 16S rRNA gene sequencing and flow cytometry, allows for monitoring the dynamics of microbial communities. Flow cytometry and amplicon sequencing fingerprinting are in agreement in terms of communities' stabilization. However, flow

cytometry's sensitivity does not allow comparison of communities presenting different species composition. Flow cytometry fingerprint can be viewed as a routine method to assess stability/disturbance of microbiota during a bioprocess. This is a promising high-throughput, low-cost, fast and operator-independent technology for microbial resource management.

Acknowledgements

References

1. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J Mol Biol 215, 403–10.
2. Barlaz, M.A., Ham, R.K., Schaefer, D.M., 1989. Mass-Balance Analysis of Anaerobically Decomposed Refuse. J. Environ. Eng. 115, 1088–1102.
3. Benbelkacem, H., Bayard, R., Abdelhay, A., Zhang, Y., Gourdon, R., 2010. Effect of leachate injection modes on municipal solid waste degradation in anaerobic bioreactor. Bioresour. Technol. 101, 5206–5212.
4. Bouchez, T., Patureau, D., Dabert, P., Wagner, M., Delgenès, J.P., Moletta, R., 2000. Successful and unsuccessful bioaugmentation experiments monitored by fluorescent in situ hybridization. Water Sci. Technol. 41, 61–68.
5. Brognaux, A., Han, S., Sørensen, S.J., Lebeau, F., Thonart, P., Delvigne, F., 2013. A low-cost, multiplexable, automated flow cytometry procedure for the characterization of microbial stress dynamics in bioreactors. Microb. Cell Factories 12, 100.
6. Delcenserie, V., Taminiau, B., Delhalle, L., Nezer, C., Doyen, P., Crevecoeur, S., Roussey, D., Korsak, N., Daube, G., 2014. Microbiota characterization of a Belgian protected designation of origin cheese, Herve cheese, using metagenomic analysis. J. Dairy Sci. 97, 6046–6056.
7. Delvigne, F., Goffin, P., 2014. Microbial heterogeneity affects bioprocess robustness: Dynamic single-cell analysis contributes to understanding of microbial populations. Biotechnol. J. 9, 61–72.
8. Delvigne, F., Zune, Q., Lara, A.R., Al-Soud, W., Sørensen, S.J., 2014. Metabolic variability in bioprocessing: implications of microbial phenotypic heterogeneity. Trends Biotechnol. 32, 608–616.

9. De Roy, K., Clement, L., Thas, O., Wang, Y., Boon, N., 2012. Flow cytometry for fast microbial community fingerprinting. Water Res. 46, 907–919.

10. Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27, 2194–2200.

11. Gurijala, K.R., Sulfita, J.M., 1993. Environmental Factors influencing methanogenesis from refuse in landfill samples. Environ. Sci. Technol 1176–1181.

12. Hettiaratchi, P., Jayasinghe, P., Tay, J.H., Yadav, S., 2015. Recent advances of biomass waste to gas using landfill bioreactor technology - A review. Curr. Org. Chem. 19, 413–422.

13. He, Y.L., Ding, Y.F., Long, Y.Q., 1991. Two cellulolytic Clostridium species: Clostridium cellulosi sp. nov. and Clostridium cellulofermentans sp. nov. Int. J. Syst. Bacteriol. 41, 306–309.

14. Huang, S., Sheng, P., Zhang, H., 2012. Isolation and identification of cellulolytic bacteria from the gut of Holotrichia parallela larvae (Coleoptera: Scarabaeidae). Int. J. Mol. Sci. 13, 2563–2577.

15. Hui, W., Jiajia, L., Yucai, L., Peng, G., Xiaofen, W., Kazuhiro, M., Zongjun, C., 2013. Bioconversion of un-pretreated lignocellulosic materials by a microbial consortium XDC-2. Bioresour. Technol. 136, 481–487.

16. Kinet, R., Destain, J., Hiligsmann, S., Thonart, P., Delhalle, L., Taminiau, B., Daube, G., Delvigne, F., 2015. Thermophilic and cellulolytic consortium isolated from composting plants improves anaerobic digestion of cellulosic biomass: Toward a microbial resource management approach. Bioresour. Technol. 189, 138–144.

17. Koch, C., Fetzer, I., Schmidt, T., Harms, H., Müller, S., 2013. Monitoring Functions in Managed Microbial Systems by Cytometric Bar Coding. Environ. Sci. Technol. 47, 1753–1760.

18. Koch, C., Gunther, S., Desta, A.F., Hubschmann, T., Muller, S., 2013. Cytometric fingerprinting for analyzing microbial intracommunity structure variation and identifying subcommunity function. Nat Protoc 8, 190–202.

19. Koch, C., Harnisch, F., Schröder, U., Müller, S., 2014. Cytometric fingerprints: evaluation of new tools for analyzing microbial community dynamics. Front. Microbiol. 5, 273.

20. Li, T., Mazéas, 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.

21. Liu, Y., Qiu, Z.P., Wang, G.C., 2013. Acceleration of the degradation of leachate with introduced microbes in bioreactor sanitary landfill. Adv. Mater. Res. 750-752, 1304–1308.

22. Marzorati, M., Wittebolle, L., Boon, N., Daffonchio, D., Verstraete, W., 2008. How to get more out of molecular fingerprints: practical tools for microbial ecology. Env. Microbiol 10, 1571–81.

23. Müller, S., Hübschmann, T., Kleinsteuber, S., Vogt, C., 2012. High resolution single cell analytics to follow microbial community dynamics in anaerobic ecosystems. Methods, Flow Cytometry and Cell Sorting: the Next Generation 57, 338–349.

24. Neumann, L., Scherer, P., 2011. Impact of bioaugmentation by compost on the performance and ecology of an anaerobic digester fed with energy crops. Bioresour. Technol. 102, 2931–2935.

25. O-Thong, S., Prasertsan, P., Karakashev, D., Angelidaki, I., 2008. Thermophilic fermentative hydrogen production by the newly isolated Thermoanaerobacterium thermosaccharolyticum PSU-2. Int. J. Hydrog. Energy 33, 1204–1214.

26. Partanen, P., Hultman, J., Paulin, L., Auvinen, P., Romantschuk, M., 2010. Bacterial diversity at different stages of the composting process. BMC Microbiol. 10, 94.

27. Peng, X., Börner, R.A., Nges, I.A., Liu, J., 2014. Impact of bioaugmentation on biochemical methane potential for wheat straw with addition of Clostridium cellulolyticum. Bioresour. Technol. 152, 567–571.

28. Pohland, F.G., Kang, S.J., 1975. Sanitary Landfill Stabilization with leachate recycle and residual treatment, in: AICheE Symp Series 145-71.

29. Reinhart, D.R., McCreanor, P.T., Townsend, T., 2002. The bioreactor landfill: Its status and future. Waste Manag. Res. 20, 172–186.

30. Ren, N., Cao, G., Wang, A., Lee, D.-J., Guo, W., Zhu, Y., 2008. Dark fermentation of xylose and glucose mix using isolated Thermoanaerobacterium thermosaccharolyticum W16. Int. J. Hydrog. Energy 33, 6124–6132.

31. Rodriguez, C., Taminiau, B., Brévers, B., Avesani, V., Van Broeck, J., Leroux, A., Gallot, M., Bruwier, A., Amory, H., Delmée, M., Daube, G., 2015. Faecal microbiota characterisation of horses using 16 rdna barcoded pyrosequencing, and carriage rate of clostridium difficile at hospital admission. BMC Microbiol. 15.

32. Rogers, W.T., Holyst, H.A., 2009. FlowFP: A Bioconductor Package for Fingerprinting Flow Cytometric Data. Adv. Bioinforma. 2009, e193947.

33. Rogers, W.T., Moser, A.R., Holyst, H.A., Bantly, A., Mohler III, E.R., Scangas, G., Moore, J.S., 2008. Cytometric fingerprinting: Quantitative characterization of multivariate distributions. Cytometry A 73, 430–441.

34. Scherer, P., Neumann, L., 2013. "Methano-compost", a booster and restoring agent for thermophilic anaerobic digestion of energy crops. Biomass Bioenergy 56, 471–478.

35. Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., Weber, C.F., 2009. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Appl. Environ. Microbiol. 75, 7537–7541.

36. Sethi, S., Kothiyal, N.C., Nema, A.K., 2013. Stabilisation of municipal solid waste in bioreactor landfills - An overview. Int. J. Environ. Pollut. 51, 57–78.

37. Vanwonterghem, I., Jensen, P.D., Ho, D.P., Batstone, D.J., Tyson, G.W., 2014. Linking microbial community structure, interactions and function in anaerobic digesters using new molecular techniques. Curr. Opin. Biotechnol. 27, 55–64.

38. Vavilin, V.A., Fernandez, B., Palatsi, J., Flotats, X., 2008. Hydrolysis kinetics in anaerobic degradation of particulate organic material: An overview. Waste Manag. 28, 939–951.

39. Wang, Y.-S., Byrd, C.S., Barlaz, M.A., 1994. Anaerobic biodegradability of cellulose and hemicellulose in excavated refuse samples using a biochemical methane potential assay. J Ind Microbiol. 13, 147–153.
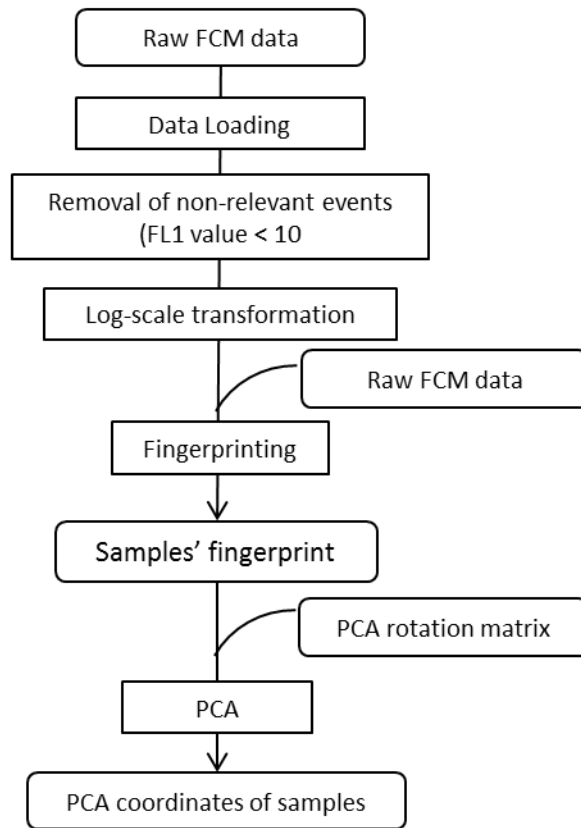
Appendix

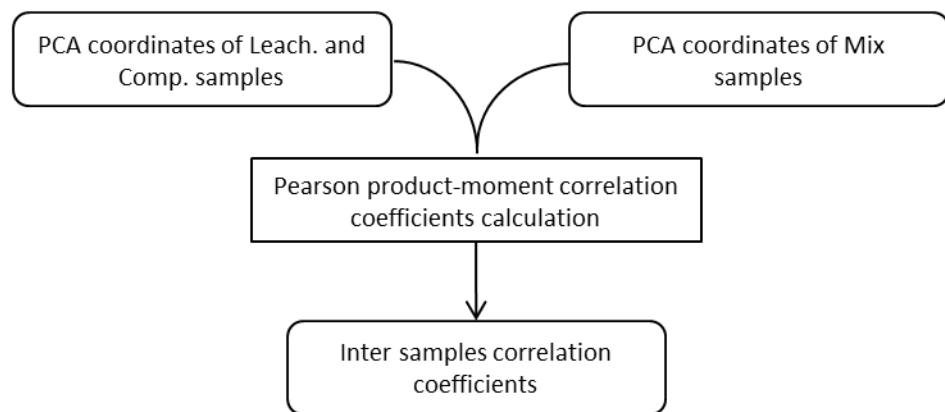Supplementary file 1 : Flow cytometric data analysis pipeline

Data treatment

Data

Data analysis pipeline for samples of compost and leachate consortium
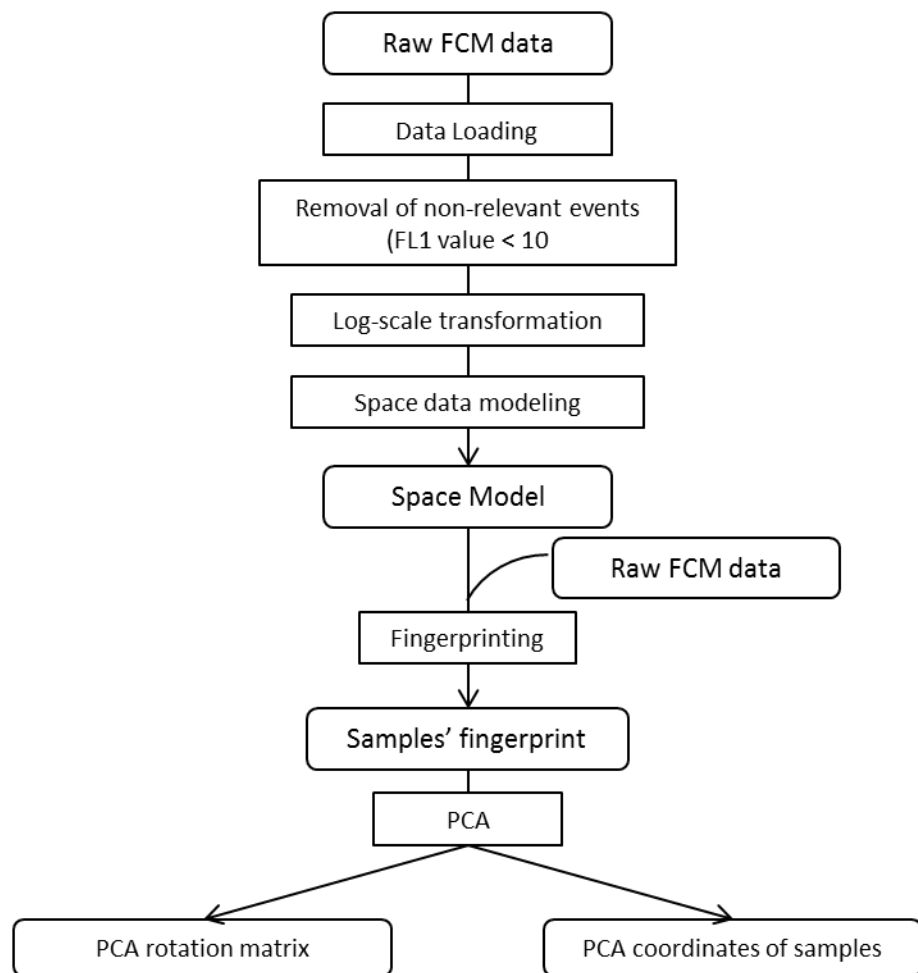
Data analysis pipeline for samples
of mix consortium

Inter samples comparison

CHAPTER III:

FLOW CYTOMETRY FOR METABOLIC FINGERPRINTING OF

SINGLE SPECIES BIOFILM

This chapter corresponds to the article entitled "*Dynamic of biofilm formation by Bacillus amyloliquefaciens deciphered by flow cytometry phenotypic fingerprinting*" (**Kinet Romain**, M. Ongena, N. Boon, F. Delvigne) which will be submitted in Microbial Biotechnology.

Abstract

Potential of biofilm has recently been pointed out for fine chemical/biological synthesis. Especially, single species biofilm hold more and more attention as they display high potentialities for biocatalytic processes. Such communities exploit stochastic gene expression (cells switching randomly between discrete physiological states by differentiating gene expression) in order to divide labor required for biofilm formation. Cooperative interactions of specialized sub-populations yield ecological benefit for the community. Therefore, biofilm microbiota cannot be considered as a 'blackbox'. A phenotypic fingerprint, representative of the isogenic community in its entirety, should be available for the efficient management of a biofilm reactor. Flow cytometry fingerprinting approach combined with RSG staining has been implemented in this work for obtaining metabolic fingerprints of biofilm communities formed by diverse *B. amyloliquefaciens* strains. This combined approach highlights various metabolic behaviors among monitored biofilm communities. Biofilm communities which are able to synthetize three typical *B. amyloliquefaciens* lipopeptides families (i.e. surfactin, fengycin, iturin) display divergent behavior from deleted communities exhibiting partial lipopeptide profile. Complementary, macroscale observation of biofilm communities reveals impacts of the alteration of native set of lipopeptide on biofilm structure. As for metabolic activities, communities with an altered lipopeptide profile are characterized by a divergent behavior by comparison with communities exhibiting the full set of lipopeptide. The maintenance of biofilm structure is negatively impacted by disequilibrium in lipopeptide profile.

Key-words

1 Introduction

Since decade, bacterial biofilms are involved in many environmental applications. Comparatively to planktonic cells, biofilm organization improves environmental continuous processes by increasing biomass concentration and its hydraulic retention time (Nicolella et al., 2000). By preventing biomass washout, biofilm formation on solid carrier results in a continuous conversion of feed stream. Traditionally, the management of such processes relies on control of few biotic factors such as feeding rate, aerobic/anaerobic conditions and shearing forces (Zune, 2015). As biofilm formation and maintenance naturally occurs in processes conditions, elucidating complex biofilm formation and maintenance is not necessary for efficiently managing such low added value applications.

However, potential of biofilm has recently been pointed out for fine chemical/biological synthesis (Cheng et al., 2010). Especially, single species biofilm hold more and more attention as they display high potentialities for biocatalytic processes (Halan et al., 2012; Rosche et al., 2009). Functional specialization and stratification occurring in monospecies biofilm improve the robustness and the productivity of processes when comparing with planktonic cell strategy. These multicellular communities exploit phenotypic heterogeneity through a division of labor strategy (Martins and Locke, 2015). More precisely, bacteria exploit stochastic gene expression (cells switching randomly between discrete physiological states by differentiating gene expression) for distributing the metabolic burden associated with biofilm formation, leading to specialized subpopulations as it has been demonstrated in *B. subtilis* biofilm (Vlamakis et al., 2008). Only a small proportion of cells has been evidenced for expressing gene encoding matrix components (Chai et al., 2008). The energy cost, inherent to the synthesis of these molecules, is saved by non-producing cells which benefit

from matrix components produced by others. Similarly, biofilm maintenance and community survival/dispersion is ensured through functional specialization in mature communities. A part of the community maintain active growth while others differentiate into several phenotypes such as matrix producer, spore former or competent cell through complex regulatory pathways (Stewart and Franklin, 2008). More than co-existence of specialized groups, division of labor strategy imply positive interactions between different groups. A colony formed of multiple phenotypes is assumed preforming better than a colony that consists solely of any one of them. Cooperative interactions of specialized sub-populations yield ecological benefit for the community.

Biofilm microbiota cannot be anymore considered as a 'blackbox'. Spatial and functional heterogeneities affecting biofilm community should be elucidated. Moreover, physiological heterogeneities and related underlying molecular mechanisms involved in biofilm formation and maintenance must be deeply understood for process optimization. Nowadays, several techniques allows for evaluation at microscale of biofilm structure and function. Particularly, fluorescent reporter systems combined with advanced imaging approach (e.g. confocal laser scanning microscopy) allow for imaging the three dimensional distribution of microorganisms within the biofilm structure (Morgenroth and Milferstedt, 2009). Flow cell which allows for a tightly control of the flow, can be easily combined with an imaging approach (Sternberg and Tolker-Nielsen, 2005). Microfluidic devices also reveal powerful for microscale investigations (Bruchmann et al., 2015; Delvigne et al., 2014). Moreover, it is crucial to link these microscale investigations with meso-/macroscale observations as mechanisms and processes at the microscale depend on and influence interactions on larger spatial scales (Morgenroth and Milferstedt, 2009).

Ideally, a phenotypic fingerprint, representative of the isogenic community in its whole, should be available for an efficient management of a biofilm reactor. Previous mentioned approaches are not suitable for this purpose. Confocal laser microscopy is a powerful tool for highlighting heterogeneities, but this technique is only able to provide information about restricted/specific targeted area of the biofilm structure. While, conditions encountered in flow cell by biofilm communities diverge from bioprocesses conditions. Based on this statement, an original flow cytometry fingerprinting approach has been implemented in this work for obtaining metabolic fingerprints of artificial biofilm communities. Metabolic activity of every cell is evaluated thanks to Redox Sensor Green (RSG) stain. Kalyuzhnaya et al. (2008) demonstrated that RSG does not suppress cellular metabolism contrary to other metabolic marker (tetrazolium salts) and is useful for real-time detection of cells actively respiring. Moreover, Baert et al. (2016) demonstrate in their works that RSG fluorescent signal is proportional to the activity of the electron transport chain. Bacterial reductases, being part of electron transport systems, induce RSG yields green fluorescence.

2 Material and methods

2.1 Microbial strains and culture conditions

Bacillus FZB42 strain and its mutant AK1, AK2, AK3, CH1 and CH2 used in this work were kindly provided by R. Borriss of Humboldt University, Berlin. Mutants were constructed in order to obtain strains with deficient in lipopeptide synthesis (Koumoutsi et al., 2004). Table 1 lists respective lipopeptide synthesis profiles of the different strains.

Table 1 Respective lipopeptide synthesis profile of FZB42 strain and its mutant

	Bacillomycine	Fengycine	Surfactine
FZB	+	+	+
AK1	-	+	+
AK2	+	-	+
AK3	-	-	+
CH1	+	+	-
CH2	+	-	-

All cultures carried out in this work are led in not shaken 24 wells plates (4 raw of 6 wells) at 30°C and with medium optimized for lipopetides production (Opt medium) (Jacques et al., 1999). Different wells are filled with 1 ml of Opt medium and subsequently inoculated with 20 µl of preculture. Precultures are led in the same conditions than plate cultures except they are carried out in 250 ml flasks. Cellular concentration in the different precultures are controlled and adjusted if necessary for ensuring equal inoculation load. In the case of co-cultures, cell concentration ratio 1:1 is carried out for inoculation.

2.2   Cultures characterization

For each sampling time, segregated analyses of biofilm and liquid phase are realized. Consequently, separation of the two phases must be achieved before further parameters measurement. First, the liquid phase is collected. The biofilm formed at liquid-air interface is delicately pierced and the undersurface liquid is collected with precaution. Biofilm is then suspended in filtered (0.2 µm) Phosphate Buffered Saline solution (8.0 g.l^{-1} NaCl, 0.2 g.l^{-1} KCl, 1.44 g.l^{-1} Na$_2$HPO$_4$, 0.24 g.l^{-1} KH$_2$PO$_4$) and collected. Sonication step is applied to samples in order to deconstruct extrapolymeric substances matrix in which cells are tangled up. Once

the cell suspensions are obtained for both phases, respective optical densities are measured at 600 nm.

2.2.1 Sporulation

Sporulation level is estimated through acridine orange (AO) staining technique involving flow cytometry for fluorescence measurement. 100 µl of cell suspension is added with 500 µl pH 3 buffered solution 1 (13.23 mM citric acid, 6.85 mM Na_2HPO_4, 0.1 mM EDTA, 0.2 M anhydrous sucrose) and 500 µl of pH 3.8 buffered solution 2 (4.95 mM citric acid, 5.5 mM Na_2HPO_4, 0.147 M NaCl) in which AO stock solution is formerly diluted 100 times. Cell suspension concentration is previously adjusted in order to reach approximate final concentration of 10^5-10^6 cells per ml. After incubation at room temperature for 10 min, samples are centrifuged and resuspended in filtered PBS. Next, fluorescence profile of the different samples is evaluated thanks to Accuri C6 flow cytometer (BD Biosciences, NJ USA). Parameters are recorded for 40 000 events with flow rate of 14 $µl.min^{-1}$ and FSC-H based threshold of 60 000 AU. Traditional FSC-H threshold (80 000 AU) is lowered in order to take into account the spore, smaller than vegetative cells.

2.2.2 Metabolic state characterization – Redox Sensor Green (RSG) staining

Metabolic state is characterized through RSG staining technique involving flow cytometry for fluorescence measurement. 1 ml of cell suspension is added with 1 µL Redox Sensor Green (Invitrogen, UK). Cell suspension concentration is previously adjusted in order to reach approximate final concentration of 10^5-10^6 cells per ml. After incubation at room temperature for 25 min, samples are centrifuged and resuspended in filtered PBS pH 7. Next, fluorescence profile of the different samples is evaluated thanks to Accuri C6 flow cytometer

(BD Biosciences, NJ USA) on the FL1 channel. Parameters are recorded for 40 000 events with flow rate of 14 µl.min^{-1} and FSC-H based threshold of 60 000 AU.

2.2.3   Lipopeptide analysis

Samples were analyzed by reverse phase UPLC–MS (UPLC, Waters, Acquity class H) coupled with a single quadrupole MS (SQDetector, Waters, Acquity) on an Acquity UPLC BEH C18 (Waters) 2.1 × 50 mm, 1.7 µm column. We used a method, based on acetonitrile gradients, allowing the simultaneous detection of all three lipopeptide families. Elution was started at 30% acetonitrile (flow rate of 0.60 ml min−1). After 2.43 min, the percentage of acetonitrile was brought up to 95% and held until 5.2 min. Then, the column was stabilized at an acetonitrile percentage of 30% for 1.7 min. Compounds were identified on the basis of their retention times compared with authentic standards (98% purity, Lipofabrik society, Villeneuve d'Asc, France) and the masses detected in the SQDetector. Ionization and source conditions were set as follows: source temperature, 130°C; desolvation temperature, 400°C; nitrogen flow, 1000 l h−1 ; cone voltage, 120 V.

2.3   Statistical data treatment

2.3.1   Flow cytometry data

The raw data are extracted from the CFlow software (Accuri, BD Bioscience) as .fcs files and loaded into R software. As well for RSG staining as for AO staining, FSC-H, SSC-H features and induced green fluorescence (FL1 channel) are considered for further data treatment. Data analysis pipeline consists of three major steps. (1) Creating a n-dimensional quantitative fingerprint of each sample from their respective bivariate FCM distributions, either (SSC − FL1) or (FSC − FL1), using the recursive probability binning (PB) algorithm for flow cytometry data, implemented in the Bioconductor package FlowFP (Rogers and Holyst, 2009). In a first

time, a model, composed of hyper-rectangular regions (bins) of varying size and shape, is established. From superimposed FCM distribution (i.e. data from all samples are pooled together) and thanks to PB algorithm (Rogers et al., 2008), bivariate data space is divided in hyper-rectangular regions in such way that each contains similar number of events (one event corresponding to one cell). First step of the algorithm consists in the division of space into two bins containing similar number of events. Afterwards, each of these bins is again divided into two bins with equal number of events, and so forth. Therefore region of bivariate FCM space displaying high density of events are characterized by bins of small area whereas larger bins characterized regions of weak density. Moreover, the final number of bins (n) is arbitrary set and correspond to 2^i with i the number of recursive subdivisions. In our case, 6 recursive subdivisions are applied and therefore model is composed of 64 bins. The obtained model is then applied to each sample and number of cells per bin is determined, creating feature vector of counts (n-dimensional) for each sample (figure 1). The latter will be used as a fingerprint of the microbial community at a given time and under given conditions. (2) For each type of strains or cocultures, a vector of values is created by computing successive feature vectors of counts obtained at step 1 for the different sampling times. (3) Hierarchical clustering of the different strains according to their respective vector of values created at step 2. Hierarchical clustering on principal component is carried out for establishing clusters.
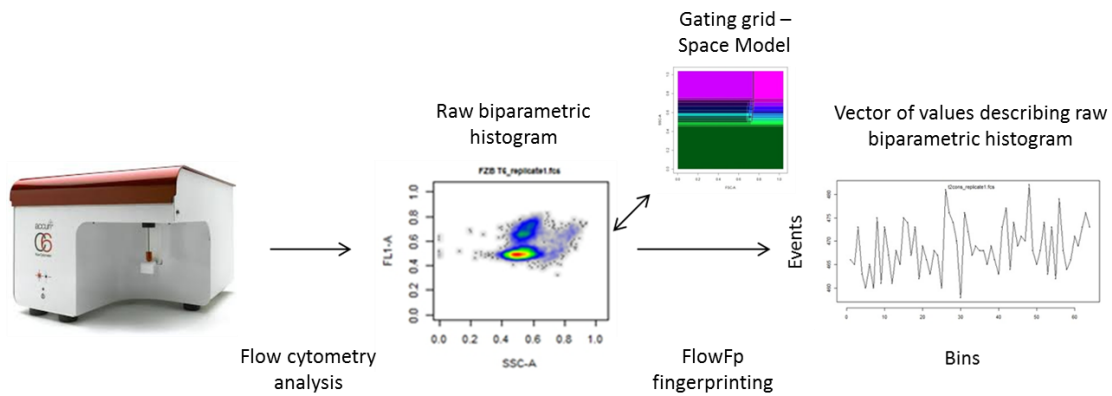
Figure 1 Flow sheet of flow cytometry fingerprinting (FlowFP). A vector of values describing raw biparametric histogram (SSC-A – FLA-1) is obtained thanks to FlowFP fingerprinting. Gating grid composed of 64 bins is applied to histograms. The number of events per bins is determined for constructing vector of values. The structure of the gating grid depends on the structure of the raw biparametric cytometric histograms studied.

3 Results and discussion

3.1 Growth dynamics of FZB42 and its mutant

B. amyloliquefaciens FZB42 and its mutants, deficient in lipopeptide synthesis, have been cultured in static conditions in order to highlight potential impact of deficient lipopeptide synthesis on biofilm formation. Additionaly, two co-cultures have been carried out, i.e. AK1-CH2 and AK3-CH1. Mixing these deletion mutants allows for the production of the full set of lipopeptides (similar to wild FZB42 strain). The ability of the different strains in forming biofilm is depicted at figure 2A through the OD values measured for the biofilm phase during culture. Moreover, figure 2B displays complementary OD values measured in liquid phase of the cultures and figure 2C displays the total cellular densities (i.e. addition of cellular densities from both liquid and biofilm phase). All strains display similar growth dynamics until 56 hours of culture, i.e. exponential growth followed by maintenance period. As well the FZB42 strain as its mutants and the co-cultures is characterized by fast exponential growth phase resulting in the fast formation of biofilm. Maximal cellular densities in biofilm

phase are reached for every strain after 26 or 32 hours. Nevertheless, initial growth is observed in liquid phase.
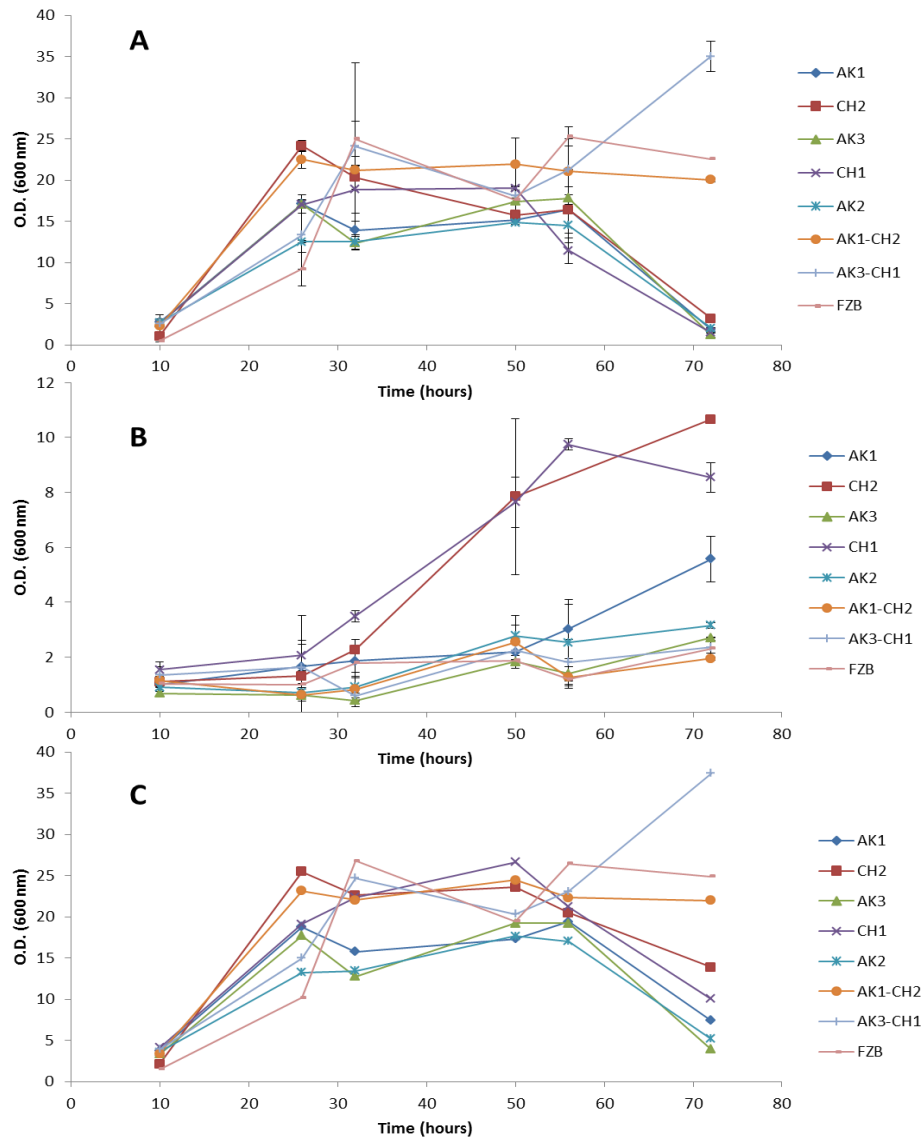

Figure 2 Dynamics of OD (600 nm) measured during not shaken culture of the different *B. amyloliquefaciens* strains: AK1 (◊), CH2 (□), AK3 (Δ), CH1 (X), AK2 (∗), AK1-CH2 (○), AK3-CH1 (+), FZB (-) in biofilm phase (A), liquid phase (B) and both (C).

An increase of OD is observed in liquid phase between inoculation and 10 hours of culture. Once critical cellular density is reached in liquid phase (e.g. an optical density of approximatively 0.5 in this experiment), cells slip towards air-liquid interface and form biofilm. Quorum sensing mechanisms are believed to trigger genetic switch responsible of

the phenotype transition (i.e. when the local density of bacteria exceeds a threshold value, the sessile phenotype is favored over motile state) (Kobayashi, 2007; Vlamakis et al., 2008). Ardré et al. (2015) also advance aerotaxis phenomenon for explaining similar fact occurring in static liquid culture of *B. subtilis*, which is closely related to *B. amyloliquefaciens*. *B. subtilis* is known to exhibit aerotaxis (i.e. bacteria migrate to areas in which the concentration of oxygen is high) and water-air interface acts as an oxygen source.

Biofilm establishment is then followed by a maintenance period. Maintenance period takes place at least until 56 hours, except for CH1 strain whose biofilm destructuration already occurs from 50th hours. From 56 hours, biofilm dispersion affects all mutant strains. After 72 hours, distinction between biofilm and liquid phase is very tough. Contrary, not any dispersion is observed for FZB42 and co-cultures. Increase of OD values in biofilm phase is even observed for AK3-CH1 co-culture.

Exopolysaccharides (EPSs), which are typically one of the main constituents of biofilm matrix, are known to be essential for the development of the architecture of any biofilm matrix (Sutherland, 2001). In this case, deficiency in some lipopetide synthesis seems to negatively impact biofilm maintenance. More than lipopeptides themselves, the interactions and equilibrium between different lipopeptide families seem crucial for stable biofilm structure. These differential biofilm behaviors, between strains exhibiting either full or partial lipopeptide profile, are corroborated by hierarchical clustering designed on the basis of respective OD dynamic profiles (figure 3). All mutants are grouped in a specific cluster whereas FZB42 and both co-culture take place in a divergent branch of the hierarchical tree.

As expected, an increase of OD values in liquid phase is observed during the dispersion of the biofilm. This phenomenon is particularly marked for AK1, CH1 and CH2 strains.

Disappearance of the biofilm organization results in the release of cells constituting biofilm to the liquid phase. The evolution of total cellular densities (figure 2C) tends towards corroborating this increase of OD in liquid phase is only the result of biofilm dispersion. The cellular densities curves measured for biofilm and both liquid-biofilm phases (total) display identical shapes. The cellular density associated to liquid phase slightly impact total cellular density. From the 32th hour no more significant increase of total OD is observed for the dispersing strains. The final increase observed in liquid phase does not result in an increase of the total population. Cells cultures enter in stationary phase after 32 hours.
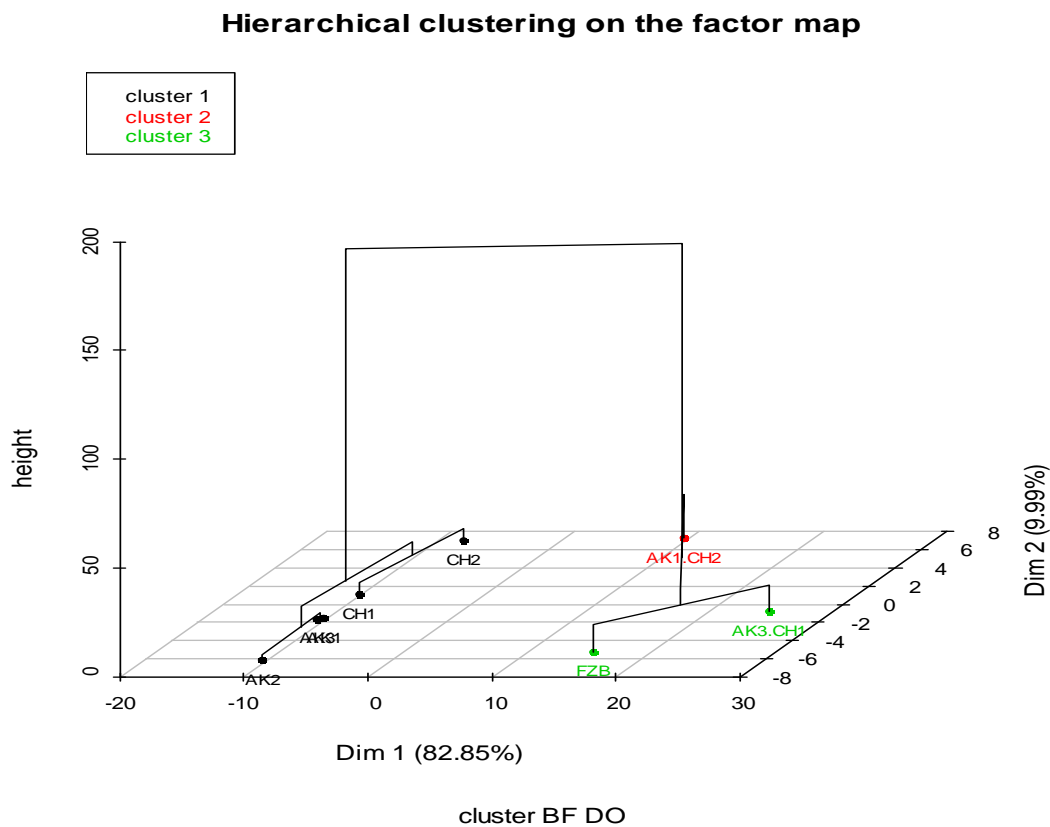

Figure 3 Hierarchical clustering based on principal component analysis of the biofilm O.D. dynamics profiles associated to different strains and co-cultures. Characteristics of the different mutants are provided in table 1.

3.2 Lipopeptide synthesis

Three types of lipopeptides are typically synthetizes by *B. amyloliquefaciens*, i.e. surfactin, fengycin and iturin. Their respective concentrations in the liquid phase have been

determined at different moment of the cultures carried out in not shaken conditions for the different strains and co-cultures. Figure 4 A, 4 B and 4 C display respectively the evolution of the concentration of surfactin, fengycin and iturin in the liquid phase.

First of all, these results confirm that the three types of lipopeptides are present in the co-cultures of mutant strains. "Native" lipopeptide profile of *B.amyloliquefaciens* is reconstructed through the co-inoculation of complementary mutant strains. Otherwise, similar dynamics of production are observed for every "strains". After 10 hours of growth, when mature biofilm structure is not yet set up, weak amounts of lipopeptide are detected. Afterwards, an important increase of the concentrations is observed for each strain between the 10th and 26th hours. Finally, surfactin and fengycin concentrations remain pretty stable until the end of the cultures while iturin concentrations still increase until the 50th hour. The three types of lipopeptide are thus mainly produced by bacteria during exponential growth during which biofilm structure is elaborated. Nevertheless, iturin and fengycin synthesis seems delayed comparatively to surfactin synthesis. In their works, Jacques et al. (1999) and Koumoutsi et al. (2004) highlighted similar dynamics in lipopetides production from respectively *Bacillus subtilis* S499 which has since been reclassified in *Bacillus amyloliquefaciens* S499 (Nihorimbere et al., 2012) and *Bacillus amyloliquefaciens* FZB42.

Except for the first sampling time, maximal lipopeptides concentrations are measured in the wild strain cultures. At first sampling time, FZB42, which display the lowest cellular density in biofilm, exhibit the lowest concentrations. Contrary, during the rest of the culture, the wild strain is characterized by the highest concentrations for the three types of lipopeptides. Particularly, fengycin concentrations are two times higher for the wild strain than for mutants and co-cultures. Moreover, these results highlight lipopeptides are not produced in the same range of concentration. Surfactin is synthetized in much lower quantities than
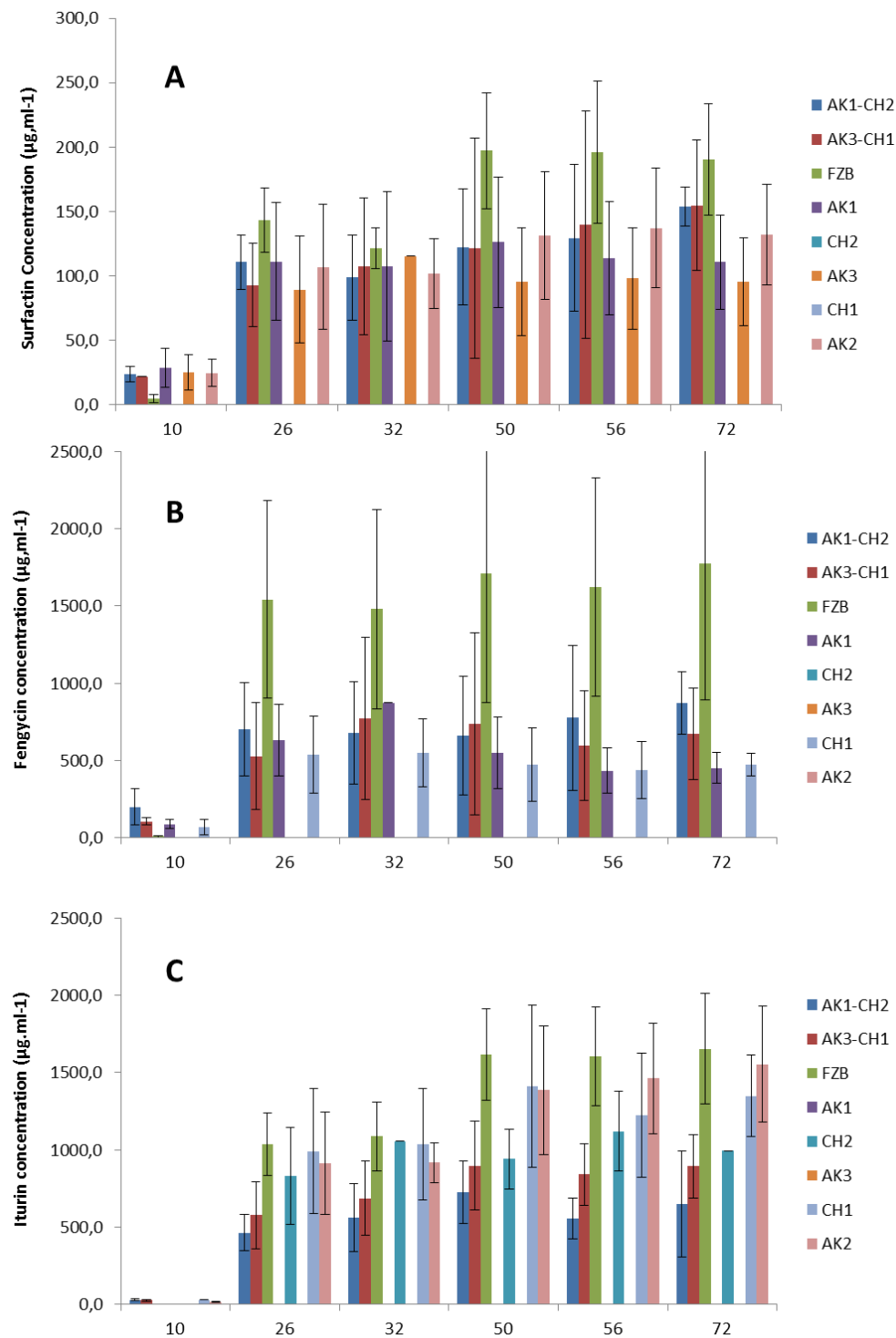
Figure 4 Lipopeptide concentrations ((A) surfactin, (B) fengycin and (C) iturin) during static microplate assays of *B.amyloliquefaciens* FZB42, mutant strains and co-cultures. Characteristics of the different mutants are provided in table 1.

fengycin and iturin. Indeed, surfactin concentrations are approximatively 10 times lower

than iturin and fengycin. These results are in contradiction with the results obtained by

Jacques et al. (1999). Surfactin was the main product (maximal concentration of ~1200 mg.L$^-$

[1]) while fengycin is produced in low quantity (~100 mg.L^{-1}). However, these productions were obtained in different condition. The microbial cultures were carried out in shaken bioreactor and not in static microplate in such way that biofilm formation was avoided. This could explain such different lipopeptide production during culture.

3.3 Sporulation dynamics

Complementary to cellular densities, sporulation ratio have been assessed for each sample in both liquid and biofilm phase (figure 5). Indeed, The segregation between vegetative and spores is based on AO staining method. Linkage of AO molecules to vegetative cells results in significantly higher fluorescent signal than linkage between AO and sporulated cells. Sporulation ratio of a microbial population can thus be determined through measurement of single cell fluorescence.

As highlighted at figure 5, spores represent a very weak proportion of the global population at the beginning of exponential growth phase (i.e. after 10 hours of culture) for every strain. However, after 24 hours, more than 60% of spores are identified in biofilm for both co-cultures and 47%±1 in FZB42's biofilm. AK1 displays intermediate sporulation ratio (32%±3) in biofilm phase after the same time whereas lower rates are calculated for other mutants (around 10%). Next, sporulation ratios increase for each type of "strains". However, unlike mutants, which reach maximal rates between 50 hours and 72 hours, FZB42 and co-cultures already reach maximal sporulation ratios after 34 hours. Moreover, while for mutants sporulation ratios remain stable after reaching maximums, a decrease is observed for the strains exhibiting a full lipopeptide profile. This decrease of the sporulation ratios between the 34th hour and the end of the process is even more pronounced in the liquid phase. Concerning liquid phase, similar trends and values than biofilm phase are observed. It is also

interesting to note that very high sporulation ratios are reached in the cultures. Indeed, maximal values oscillate between 80% and 95%. Similar ratios have been measured by Wijman et al. (2007) in biofilms formed by *Bacillus cereus*.
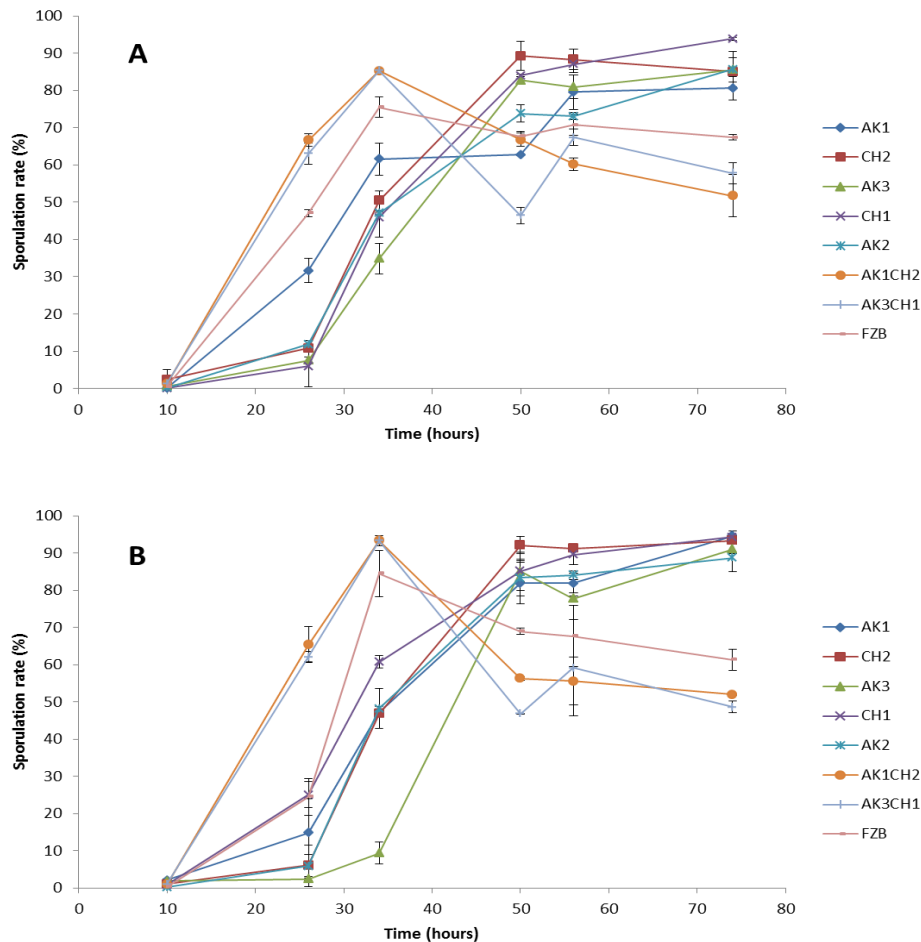

Figure 5 Dynamics of sporulation measured during static culture of the different *B. amyloliquefaciens* strains: AK1 (◊), CH2 (□), AK3 (Δ), CH1 (X), AK2 (✳), AK1-CH2 (○), AK3-CH1 (✚), FZB (-) in biofilm (A) and liquid phase (B). Sporulation ratios are determined through AO staining approach.

3.4 Metabolic activities dynamics

3.4.1 Median values

The dynamics of metabolic activities for the different strains in liquid and biofilm phase are displayed at figure 6. The metabolic activities are expressed through the median values of fluorescence resulting from RSG staining and measured for 40 000 cells for each sample by flow cytometry.

Displaying their highest level of metabolic activities after 10 hours of culture (median values comprised between 150 000 and 280 000), biofilm communities exhibit drastic lower metabolic activities after 24 hours (median values around 30 000) (figure 6 A). Afterwards, metabolic activities associated to biofilm cells still decrease for finally stabilizing at a minimal level after 50 hours of culture. Once biofilm structure is established, bacteria enter in a weakly metabolically active stage. Such dynamics is observed for every type of strain in biofilm phase. Moreover, quantitative differences only appear for the first sampling time, before the exponential growth phase.
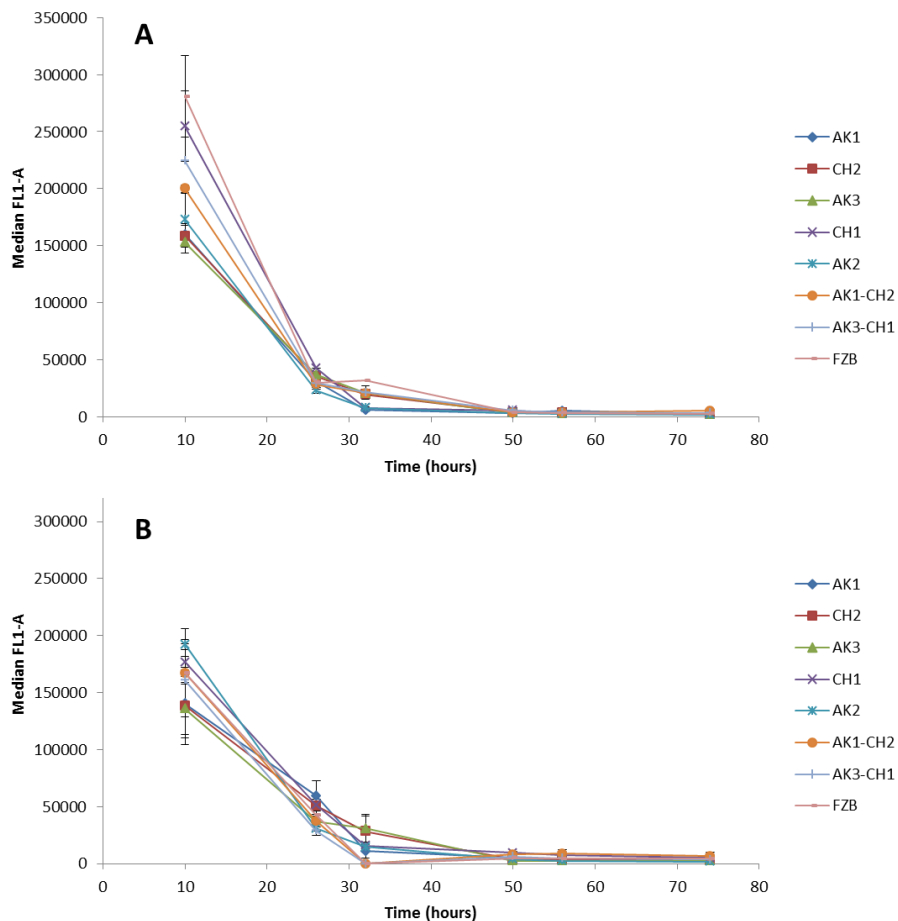

Figure 6 Dynamics of the metabolic activities measured through RSG staining during static culture of the different *B. amyloliquefaciens* strains: AK1 (◊), CH2 (□), AK3 (Δ), CH1 (X), AK2 (✳), AK1-CH2 (○), AK3-CH1 (✛), FZB (-) in biofilm (A) and liquid phase (B). Plotted values correspond to the median of FL1-A values obtained for the different samples.

Except a slower initial decrease of the metabolic activities, similar dynamics are observed for cells extracted from liquid phase and from biofilm. Such similar trend would suggest that cells observed in liquid phase result from biofilm dispersion and are not the result of specific planktonic cells growth. The decrease of metabolic activities associated to bacteria from liquid phase (figure 6 B) corroborates the hypothesis that the final increase of OD, observed in the liquid phase, is only the result of the release of cells from biofilm.

3.4.2 Metabolic phenotypic fingerprinting

Bidimensional flow cytometric representations resulting from RSG staining approach implemented for the investigation of metabolic activities of the different strains are displayed at figure 7. More precisely, this figure displays cytometric patterns (FSC-FL1 and SSC-FL1) obtained for the wild type (FZB42) biofilm community during culture.

In the previous section, median FL1-A values related to these distributions have been discussed. However, a part of the information associated to the cytometric analyses is lost by observing such discrete values. The distribution/diversity of the metabolic single-cell activities is not taking into account. Just by giving a look at the cytogram (figure 7), it is obvious that metabolically different subpopulations appear in biofilms. The dynamic of these different metabolic phenotypes is neglected when only interpreting median values. Based on this statement, metabolic phenotypic fingerprinting (see section 2.3. for data analysis pipeline) has been implemented with the aim of clustering the strains according to their metabolic behavior during cultures. Figure 8 and the figure 9 display the results of these metabolic clustering when respectively considering SSC-FL1 and FSC-FL1 bidimensional cytometric patterns related to biofilm cells. As cells from liquid phase are thought to be only

the result of biofilm release, phenotypic fingerprinting has not been considered for these ones.

Divergent behavior of FZB42 and co-cultures observed for biofilm dynamics and highlighted through O.D. clustering, is also noticed in term of metabolic phenotypic dynamics. The shape of both hierarchical trees' highlights this divergence. FSC and SSC parameters of the cells, which typically give structural information about respectively cell size and about internal cell organization, do not impact clustering. At least, FSC or SSC variability measured between cells is not significant compare to FL1 variability. As for cellular densities clustering, "strains" with an uncomplete lipopeptide profile are strongly separated from "strains" with complete set of lipopeptide. While AK1, AK2, AK3, CH1 and CH2 are grouped in the same branch of the trees, FZB42 and co-culture take place in a divergent branch. Similar divergences are highlighted at macroscale, i.e. dynamic of biofilm formation, and microscale, i.e. dynamics of metabolic phenotypes.

By comparing metabolic behaviors of co-cultures of mutant strains and wild strain FZB 42, no adverse effect of the genetic construction is highlighted. Depletion in lipopeptide is confirmed as the sole potential cause for divergent biofilm metabolic behavior. A divergent biofilm structure induced by lipopeptide depletion could result in the implementation of a divergent single-cell metabolic strategy by the biofilm community.
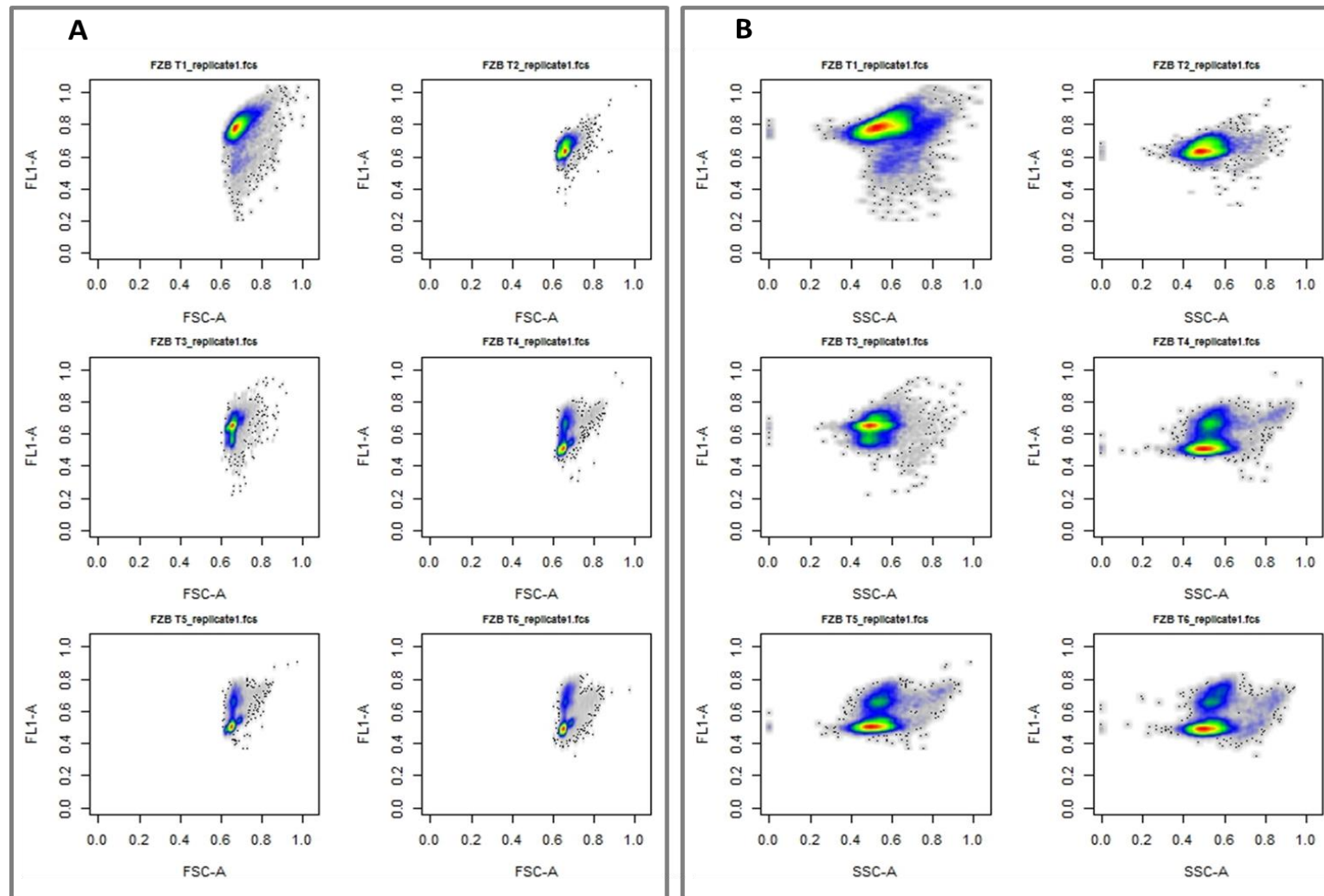
Figure 7 Bidimensional flow cytometric histogram obtained during static culture of *B. amyloliquefaciens* FZB42((A) FSC-A/FL1-A and (B) SSC-A/FL1-A). Fluorescence, expressed through FL1-A values, results from RSG staining technique. Histograms are displayed for six sampling times for both cases.
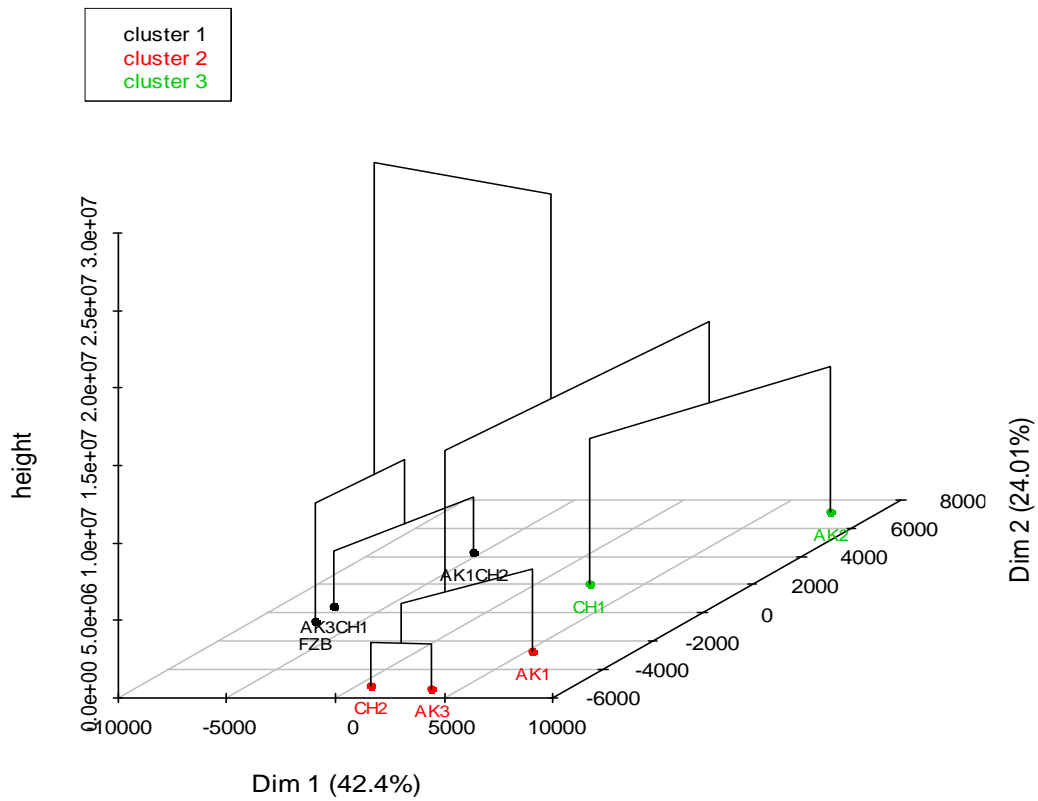
Figure 8 Hierarchical clustering based on principal component analysis of the biofilm metabolic flow cytometric fingerprinting (SSC-FL1) profiles associated to different strains and co-cultures.
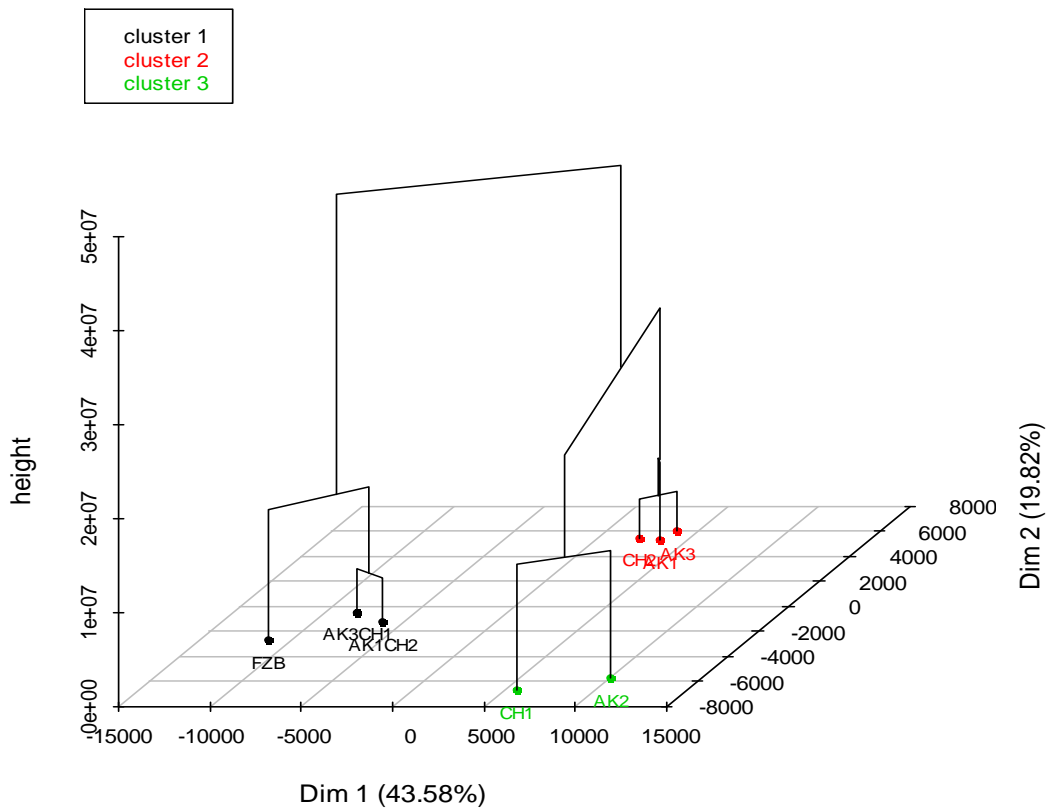

Figure 9 Hierarchical clustering based on principal component analysis of the biofilm metabolic flow cytometric fingerprinting (FSC-FL1) profiles associated to different strains and co-cultures.

III. 4 Conclusion

The flow cytometry fingerprinting approach implemented in this work has been shown as an efficient approach for the monitoring of biofilm communities at a single-cell level. Remarkably, contrary to other single-cell monitoring approach (e.g. confocal microscopy laser scanning), the whole biofilm community is characterized based on this approach. Moreover, results reveal RSG as an efficient tool for evaluating metabolic activity of cells. Combined with flow cytometry fingerprinting, RSG staining allows for highlighting various metabolic behaviors among monitored biofilm communities. Biofilm communities which are able to synthetize three typical *B. amyloliquefaciens* lipopeptides families (i.e. surfactin, fengycin, iturin) display divergent behavior from communities characterized by uncomplete lipopeptide profile. Disequilibrium in lipopeptide synthesis also impacts the architecture/structure of biofilm. The lack of one or two lipopeptide families results in a faster destabilization of the structure. Impacts of the alteration of native set of lipopeptide are highlighted for biofilm communities at both macroscale and microscale. However, further analyses are necessary for elucidating biological mechanisms responsible of these divergences and evaluating specific impact of each lipopeptide family.

References

1. Ardré, M., Henry, H., Douarche, C., Plapp, M., 2015. An individual-based model for biofilm formation at liquid surfaces. Phys. Biol. 12, 66015. doi:10.1088/1478-3975/12/6/066015

2. Baert, J., Delepierre, A., Telek, S., Fickers, P., Toye, D., Delamotte, A., Lara, A.R., Jaén, K.E., Gosset, G., Jensen, P.R., Delvigne, F., 2016. Microbial population heterogeneity versus bioreactor heterogeneity: evaluation of Redox Sensor Green as an exogenous metabolic biosensor. Eng. Life Sci. n/a-n/a. doi:10.1002/elsc.201500149

3. Bruchmann, J., Sachsenheimer, K., Rapp, B.E., Schwartz, T., 2015. Multi-channel microfluidic biosensor platform applied for online monitoring and screening of biofilm formation and activity. PloS One 10, e0117300. doi:10.1371/journal.pone.0117300

4. Chai, Y., Chu, F., Kolter, R., Losick, R., 2008. Bistability and biofilm formation in Bacillus subtilis. Mol. Microbiol. 67, 254–263. doi:10.1111/j.1365-2958.2007.06040.x

5. Cheng, K.-C., Demirci, A., Catchmark, J.M., 2010. Advances in biofilm reactors for production of value-added products. Appl. Microbiol. Biotechnol. 87, 445–456. doi:10.1007/s00253-010-2622-3

6. Delvigne, F., Zune, Q., Lara, A.R., Al-Soud, W., Sørensen, S.J., 2014. Metabolic variability in bioprocessing: implications of microbial phenotypic heterogeneity. Trends Biotechnol. 32, 608–616. doi:10.1016/j.tibtech.2014.10.002

7. Halan, B., Buehler, K., Schmid, A., 2012. Biofilms as living catalysts in continuous chemical syntheses. Trends Biotechnol. 30, 453–465. doi:10.1016/j.tibtech.2012.05.003

8. Jacques, P., Hbid, C., Destain, J., Razafindralambo, H., Paquot, M., Pauw, E.D., Thonart, P., 1999. Optimization of biosurfactant lipopeptide production from Bacillus subtilis S499 by Plackett-Burman design. Appl. Biochem. Biotechnol. 77, 223–233. doi:10.1385/ABAB:77:1-3:223

9. Kalyuzhnaya, M.G., Lidstrom, M.E., Chistoserdova, L., 2008. Real-time detection of actively metabolizing microbes by redox sensing as applied to methylotroph populations in Lake Washington. ISME J. 2, 696–706. doi:10.1038/ismej.2008.32

10. Kobayashi, K., 2007. Bacillus subtilis Pellicle Formation Proceeds through Genetically Defined Morphological Changes. J. Bacteriol. 189, 4920–4931. doi:10.1128/JB.00157-07

11. Koumoutsi, A., Chen, X.-H., Henne, A., Liesegang, H., Hitzeroth, G., Franke, P., Vater, J., Borriss, R., 2004. Structural and Functional Characterization of Gene Clusters Directing Nonribosomal Synthesis of Bioactive Cyclic Lipopeptides in Bacillus amyloliquefaciens Strain FZB42. J. Bacteriol. 186, 1084–1096. doi:10.1128/JB.186.4.1084-1096.2004

12. Martins, B.M., Locke, J.C., 2015. Microbial individuality: how single-cell heterogeneity enables population level strategies. Curr. Opin. Microbiol., Cell regulation 24, 104–112. doi:10.1016/j.mib.2015.01.003

13. Morgenroth, E., Milferstedt, K., 2009. Biofilm engineering: Linking biofilm development at different length and time scales. Rev. Environ. Sci. Biotechnol. 8, 203–208. doi:10.1007/s11157-009-9163-1

14. Nicolella, C., Loosdrecht, M.C.M. van, Heijnen, S.J., Nicolella, C., Loosdrecht, M.C.M. van, Heijnen, S.J., Nicolella, C., Loosdrecht, M.C.M. van, Heijnen, S.J., Nicolella, C.,

Loosdrecht, M.C.M. van, Heijnen, S.J., Nicolella, C., Loosdrecht, M.C.M. van, Heijnen, S.J., 2000. Particle-based biofilm reactor technology. Trends Biotechnol. 18, 312–320. doi:10.1016/S0167-7799(00)01461-X

15. Nihorimbere, V., Cawoy, H., Seyer, A., Brunelle, A., Thonart, P., Ongena, M., 2012. Impact of rhizosphere factors on cyclic lipopeptide signature from the plant beneficial strain Bacillus amyloliquefaciens S499. FEMS Microbiol. Ecol. 79, 176–191. doi:10.1111/j.1574-6941.2011.01208.x

16. Rogers, W.T., Holyst, H.A., 2009. FlowFP: A Bioconductor Package for Fingerprinting Flow Cytometric Data. Adv. Bioinforma. 2009, e193947. doi:10.1155/2009/193947

17. Rogers, W.T., Moser, A.R., Holyst, H.A., Bantly, A., Mohler III, E.R., Scangas, G., Moore, J.S., 2008. Cytometric fingerprinting: Quantitative characterization of multivariate distributions. Cytometry A 73, 430–441. doi:10.1002/cyto.a.20545

18. Rosche, B., Li, X.Z., Hauer, B., Schmid, A., Buehler, K., 2009. Microbial biofilms: a concept for industrial catalysis? Trends Biotechnol. 27, 636–643. doi:10.1016/j.tibtech.2009.08.001

19. Sternberg, C., Tolker-Nielsen, T., 2005. Growing and Analyzing Biofilms in Flow Cells, in: Current Protocols in Microbiology. John Wiley & Sons, Inc.

20. Stewart, P.S., Franklin, M.J., 2008. Physiological heterogeneity in biofilms. Nat. Rev. Microbiol. 6, 199–210. doi:10.1038/nrmicro1838

21. Sutherland, I.W., 2001. The biofilm matrix – an immobilized but dynamic microbial environment. Trends Microbiol. 9, 222–227. doi:10.1016/S0966-842X(01)02012-1

22. Vlamakis, H., Aguilar, C., Losick, R., Kolter, R., 2008. Control of cell fate by the formation of an architecturally complex bacterial community. Genes Dev. 22, 945–953. doi:10.1101/gad.1645008

23. Wijman, J.G.E., Leeuw, P.P.L.A. de, Moezelaar, R., Zwietering, M.H., Abee, T., 2007. Air-Liquid Interface Biofilms of Bacillus cereus: Formation, Sporulation, and Dispersion. Appl. Environ. Microbiol. 73, 1481–1488. doi:10.1128/AEM.01781-06

24. Zune, Q., 2015. Design of a single-species biofilm reactor based on metal structured packing for the production of high added value biomolecules (Thesis). Gembloux Agro-Bio Tech, ULg.

5 General discussion and perspectives

The global objective of the project was to improve the comprehension of microbial processes underlying anaerobic cellulose digestion for improving process performances on the basis of biological parameters. In a first time, the efficiency of a bioaugmentation treatment, for improving cellulose anaerobic digestion capacities of microbial communities, has been validated. Afterwards, 16s rRNA sequencing and flow cytometry, two culture independent methods, have been considered for the monitoring of the microbial communities involved in the cellulose anaerobic digestion processes. These techniques have been considered for elucidating the complexity inherent to microbial communities through the evaluation of several parameters linked with microbial ecology, i.e. richness, evenness and dynamics. Finally, flow cytometry monitoring technique has been transposed for the monitoring of single-species biofilm multicellular communities with the aim of highlighting phenotypical heterogeneities among such isogenic population. Major results obtained for these different topics are discussed in this section.

5.1 Improving anaerobic digestion of cellulosic biomass: toward a microbial resource management approach

Among several considered environments, compost provides the community of microorganisms displaying the highest cellulolytic potential in anaerobic and thermophilic conditions. Surprisingly, none consortium able to efficiently degrade cellulose under aerobic conditions could be isolated from compost (Wongwilaiwalin et al., 2010).

The isolated community displays similar degradation capacities than well-known anaerobic cellulolytic species Clostridium thermocellum (Tachaapaikoon et al., 2012). Regardless the

substrate's level of crystallinity, the isolated consortium is able to degrade a significant amount of cellulose in a weak lapse of time. The isolated community seems to have a complete set of enzymes acting synergistically for the efficient degradation of different forms of cellulose. This absence of negative impact from crystallinity is remarkable as this factor is regularly mentioned for explaining the recalcitrance of cellulose toward hydrolysis (Mansfield et al., 1999). On the opposite and in agreement with the works led by Salehian et al. (2013), lignin is shown for clearly negatively impacting the anaerobic digestion of cellulosic substrate. Non degradable in the absence of oxygen, lignin polymers protect cellulose from hydrolysis.

Once isolated from compost, the consortium has been considered as a bioaugmentation agent for improving performance of two microbial communities involved in anaerobic digestion bioprocesses:

(1) Sludge from full-scale anaerobic digester fed with agro-food organic wastes and agricultural wastes

(2) Leachate from landfill containing municipal solid waste.

In both cases, management of the microbial population leads to an increase of the biogas production from cellulose digestion (average gains of respectively 14% and 122%). Particularly, bioaugmentation treatment results in an average improvement of 41% of the leachates' cellulolytic potential. In this specific case, bioaugmentation treatment clearly accelerates the establishment of an efficient cellulolytic population. Based on these results and other similar works (Martin-Ryals et al., 2015; Peng et al., 2014), bioaugmentation with cellulolytic microbial agent can thus be considered as a promising approach for improving digestibility of recalcitrant cellulosic substrate during anaerobic digestion. By comparison

with physical, chemical and biological pretreatment, efficient management of the microbial resources involved in processes does not significantly increase the cost of the global process and does not generate inhibitory compounds (Palmqvist and Hahn-Hägerdal, 2000).

5.2 Amplicon sequencing for elucidating complex microbial communities

In order to improve the management of the microbial resources, it is crucial for elucidating the structure of microbial communities. As several factors make classical culture dependent analytical techniques inefficient, culture independent method (i.e. 16S rRNA gene sequencing (figure 1)) has been carried out for this purpose. Particularly, the richness, the evenness and the dynamics of the microbial communities have been evaluated through the calculation of several ecological parameters on the basis of sequencing results. In a first time, consortium isolated from compost has been characterized during cellulose anaerobic digestion. Specific culture conditions (i.e. cellulose as only substrate, strict anaerobia, thermophilic conditions …) induce a drastic decrease of the richness comparatively to raw compost population (Gladden et al., 2011). Moreover, community is also characterized by a weak evenness. Only few phylotypes display successively high relative abundance. Such structure is classically recognized as not favorable for the stability of the ecosystem. Highest evenness increases the probability that a species tolerant to perturbation is present in significant proportion (Wittebolle et al., 2009) which improves the community robustness. However, despite unfavorable structure, deep modifications of the community allow for maintaining the cellulolytic capacities in evolving conditions. Similar sequencing analyses carried out for the assessment of landfill leachate cellulolytic bioaugmentation also display the importance of the succession of microbial phylotypes in the hydrolysis of cellulose and the subsequent step leading to biogas generation. Therefore, monitoring of the populations' dynamics is crucial to ensure the efficiency of such processes. However, 16S rRNA gene

sequencing cannot be considered as a suitable approach for routine on-site analysis. Based on this statement, flow cytometry fingerprinting has been implemented as a cheaper, faster and easier communities monitoring approach. By measuring a set of optical parameters, flow cytometry allows the characterization of cellular populations at single-cell level in a very short time. Becoming widely used for the monitoring of phenotypic heterogeneities within isogenic cultures (Delvigne et al., 2014; Delvigne and Goffin, 2014), flow cytometry has also already been implemented for the investigation of microbial communities structure (De Roy et al., 2012; Günther et al., 2012).

5.3 Flow cytometry fingerprinting – an efficient link with MRM parameters

Complementary to sequencing analyses, flow cytometry fingerprinting has been performed for the assessment of leachate bioaugmentation. Through a simple experimental procedure involving DNA staining (figure 2), an operator-independent fingerprint of a community is obtained in less than one hour. Afterwards, diverse statistical analyses are implemented for an efficient interpretation of the community structure and more precisely for the evaluation of similarities between communities.

The fingerprinting method has been implemented for the monitoring of the dynamics/stability of microbial communities involved in the anaerobic digestion of cellulose. Results highlight divergent behavior of the investigated populations. Particularly, flow cytometry fingerprinting underlines the time required for the establishment of a stable structure. Leachate initial population required longer period than other for adapting to the environmental conditions. Similar statistical tools have been applied to the sequencing fingerprints for assessing the relevance of the cytometric approach. Comparison of flow cytometry and amplicon sequencing show similar trends about the population dynamics of

the respective communities. Flow cytometric hypotheses are corroborated by sequencing results. Therefore flow cytometry fingerprinting can be considered as a suitable routine approach (low-cost, fast, on-site, operator independent) for the monitoring of complex microbial communities and particularly for the assessment of structure dynamics.

In a second time, the establishment/resilience of added species in bioaugmented leachate community has been investigated through flow cytometric fingerprinting. The levels of similarity between the different communities have been evaluated. As for stability assessment, sequencing fingerprints have been considered as referential for the validation of trends drawn on the basis of cytometric results. Unfortunately and contrary to dynamics analyses, results from both techniques are not in agreement. Cytometric approach does not allow a fine discrimination of the diverse genera/species involved in communities. 16S rRNA gene sequencing remains essential for this purpose.

Based on these results, flow cytometric fingerprinting appears as an efficient first line method for the detection of disturbance among communities and 16S rRNA sequencing as a support technique for understanding and explaining such disturbances. Taken together, both techniques should be suitable for the efficient management of microbial resources involved in bioprocesses.

5.4 Extension of flow cytometry fingerprinting to metabolic fingerprinting of biofilms

In this last part of our works, focus has been put on the monitoring of microbial biofilm communities, more precisely single-species biofilm. Despite an homogeneous nucleic
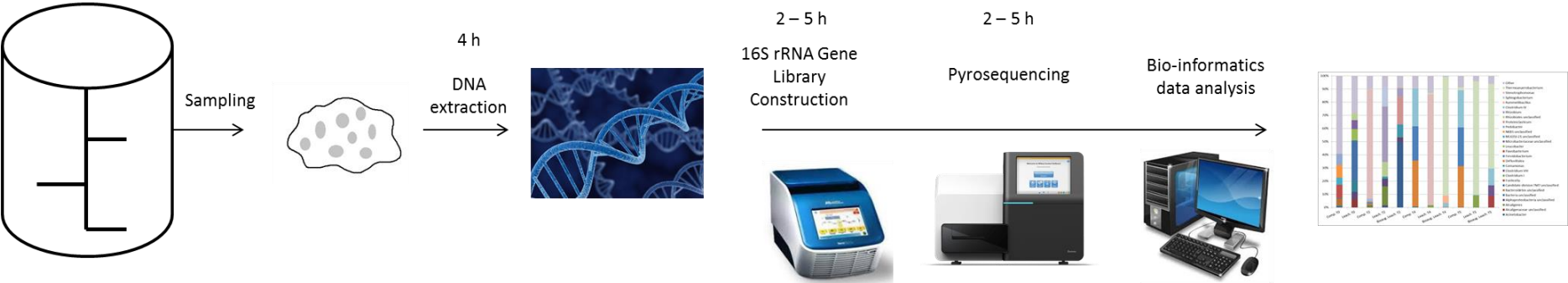
Figure 1 16S rRNA gene sequencing approach for the monitoring of complex microbial communities involved in bioprocesses
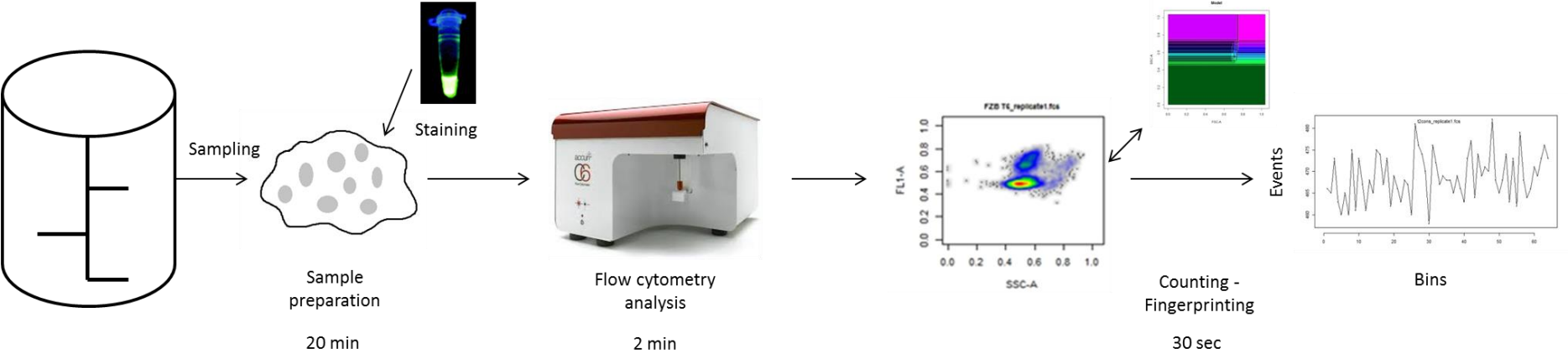

Figure 2 Flow cytometry fingerprinting approach for the monitoring of complex microbial communities involved in bioprocess

content, such multicellular communities display phenotypic/metabolic heterogeneities which result from diverse factors (e.g. division of labor strategy (Martins and Locke, 2015) and heterogeneous environmental conditions (Stewart and Franklin, 2008)). Elucidating the microbial complexity of such "multiphenotypes" communities is required for the efficient management/design of the bioprocesses which they are involved in. Physico-chemical monitoring, actually implemented is not sufficient for this purpose. A microbial single-cell approach is required. Moreover, for being truly efficient, the investigating approach should consider community in its entirety and provide information representative of the bioprocesses conditions.

Demonstrated as an efficient way to parametrize the dynamics of complex communities displaying several subpopulations (Kinet et al., 2016), flow cytometry fingerprinting has been considered for the monitoring of the heterogeneities exhibited by biofilm communities. Specifically, metabolic heterogeneities have been investigated thanks to flow cytometry fingerprinting combined with a staining technique. Redox Sensor Green, whose fluorescence is proportional to the flux of electron involved in the respiratory chain (Baert et al., 2016), is suitable for this purpose. The staining of biofilm communities, formed by *Bacillus amyloliquefaciens,* results in the discrimination of several subcommunities displaying divergent level of metabolic activities.

In addition to this microscale investigation (i.e. single-cell microbial community), biofilms has been also investigated from macroscale point of view (i.e. global biofilm structure) with the aim of highlighting potential concordances. More precisely, the impact of lipopeptide profiles on the biofilm structure formation (macroscale observations) and on the metabolic strategy implemented by biofilm communities at single-cell level (microscale observations)

have been explored. Several mutant strains of *Bacillus amyloliquefaciens*, for which the synthesis ability of one or two type of lipopeptides (i.e. surfactine, iturine, fengycine) is depleted (Koumoutsi et al., 2004), have been considered for this purpose. Divergent behaviors are observed for the biofilm communities according to their lipopeptide profile. The destabilization of the "natural" lipopeptide equilibrium results in an altered biofilm structure. Similarly to exopolysaccharides (Sutherland, 2001), lipoppeptides seem to play a key role in biofilm structure. As well from a macroscale point of view than from single-cell point of view, communities with a complete set of lipopeptide diverge from communities displaying an uncomplete profile. The maintenance of the biofilm structure is altered by a modification of the "wild" lipopeptide profile. The lack of at least one type of lipopeptide induces a faster dispersion of the biofilm structure. Similarly, a divergent metabolic behavior is highlighted for the communities characterized by an uncomplete lipopetide profile. Highlighting such divergence at single-cell level is enabled by the implementation of the fingerprinting approach for the analysis of the flow cytometry results. Fingerprinting approach parametrizes the dynamics of subcommunities revealed by staining during cultures. Contrary, the interpretation of the fluorescence (resulting from metabolic staining), through discrete values (median or mean), results in the loss of such dynamic information. Given to their respective fingerprints, strains have been clustered. While mutant strains are grouped together, communities with a complete lipopeptide profile (i.e. wild strain and co-culture of complementary mutants) take place in a divergent branch.

Previously proved to be an efficient tool for assessing the stability of multispecies communities (Kinet et al., 2016), the results obtained in this last part of our works highlight that flow cytometry fingerprinting is also relevant for the monitoring of phenotypical heterogeneities among multicellular isogenic communities such as biofilms. Moreover, the

146

approach implemented in this section allows for overcoming drawbacks inherent to methods classically used for elucidating biofilm complexity: (1) the behavior of the whole community is simultaneously investigated and (2) it is compatible with an implementation in bioprocess conditions. Nevertheless, a major drawback of the flow cytometry fingerprinting approach cannot be ignored: no biological meanings can be associated to the fingerprints obtained for the different communities.

5.5   General Conclusion

The management of microbial resources has been proved to be an efficient tool for overcoming drawbacks of bioprocesses. Particularly, bioaugmentation treatment with a cellulolytic agent (consortium) improves the efficiency of microbial communities for producing biogas from anaerobic digestion of (ligno-)cellulosic substrates. By complementing endogenous communities with a consortium isolated from compost and displaying high cellulolytic abilities, the limitation induced by the hydrolysis of complex cellulosic substrate is minimized. However, in order to improve their management, the complexity of microbial communities must be elucidated (Verstraete et al., 2007). 16S rRNA gene sequencing is a powerful tool for elucidating the structure and behavior of complex microbial communities such as those involved in cellulose anaerobic digestion. It allows for identifying major actors of the processes and for evaluating their relative distribution. Moreover, it allows for the monitoring of the dynamics of the species inside the communities. Based on sequencing results, several parameters describing the community in these terms can be calculated (richness, evenness, dynamics, …). Nevertheless, such approach is not compatible with a routine on-site implementation. Flow cytometry fingerprinting of the communities has been shown as a promising approach for routine assessment of communities' stability/disturbances in bioprocesses conditions. Combined with the staining of cellular

nucleic content, flow cytometry fingerprinting is validated as an efficient tool for the monitoring of communities dynamics.

Similarly to multispecies communities, the complexity of isogenic biofilm communities must be resolved for an efficient management of processes in which they are involved. Based on this statement and on the encouraging results previously obtained (Kinet et al., 2016), flow cytometry fingerprinting has been implemented in combination with a staining technique for the monitoring of metabolic heterogeneities among multicellular biofilm communities. Staining with Redox Sensor Green reveals that isogenic biofilm communities are composed of several subpopulations displaying divergent level of metabolic activities. Moreover, flow cytometry fingerprinting carried out for analyzing the dynamic behavior of these subcommunities during static cultures allows for highlighting similarity/dissimilarity between several investigated strains. Specifically, the impact of the alteration of the lipopeptides set on the metabolic behavior of biofilm communities is evidenced through this combined approach. The flow cytometry fingerprinting approach presented in this work allows for investigating the dynamic behavior of biofilm communities in their entirety at a single-cell level and in bioprocess conditions.

5.6 Perspectives

The results obtained during this project point out that an efficient management of microbial resources is beneficial for bioprocesses performances. In this way, flow cytometry fingerprinting proved to be an efficient and promising tool for the routine monitoring of complex microbial populations in bioprocess conditions. However, despite promising results obtained, several drawbacks are still inherent to the different methods implemented. These perspectives propose several ways for overcoming these limitations.

As shown in the second part of our works, the staining technique implemented in combination with flow cytometry fingerprinting for the monitoring of multispecies communities does not allow for discriminating members at species or even genus level. A more discriminant technique is required for identifying different members among a community. DAPI (4',6-diamidino-2-phenylindole) molecule, another fluorescent staining molecule, should be considered for this purpose (Koch et al., 2013c). Contrary to the stain involved in our works which stains all DNA without distinction, DAPI is a molecule which preferentially binds to DNA regions rich in adenine and thymine. This property results in a discrimination of the cells based on their A/T content and not just any more on their total DNA content. A FISH (fluorescence in situ hybridization) approach is also an interesting alternative for an efficient monitoring of the members of a complex community (Neumann and Scherer, 2011; Scherer and Neumann, 2013). After their identification major actors of the community could be monitored during processes through the use of fluorescent probes. However, the high specificity of this technique is also a drawback as only few actors (considered as most important) of the communities would be considered. Minor actors, which nevertheless can reveal essential for the global behavior of the community, would be neglected. Moreover, preliminary identification of the actors is imperative for the design of the fluorescent probes.

Another drawback of the flow cytometry fingerprinting technique implemented in our work is the lack of biological meaning associated to the fingerprints. The information resulting from staining are lost when fingerprints are designed. "Reference fingerprints" corresponding to the potential biological states encountered by community during bioprocess should be preliminary available. Taking into account this referential, biological meaning could be associated to the fingerprints and the evolution of the fingerprints could

be biologically explained. Similarly, the evolution of the fingerprints could be related to the

evolution of the environmental parameters as already proposed by Koch et al. (2013).

5.7 References

1. Baert, J., Delepierre, A., Telek, S., Fickers, P., Toye, D., Delamotte, A., Lara, A.R., Jaén, K.E., Gosset, G., Jensen, P.R., Delvigne, F., 2016. Microbial population heterogeneity versus bioreactor heterogeneity: evaluation of Redox Sensor Green as an exogenous metabolic biosensor. Eng. Life Sci. n/a-n/a. doi:10.1002/elsc.201500149

2. De Roy, K., Clement, L., Thas, O., Wang, Y., Boon, N., 2012. Flow cytometry for fast microbial community fingerprinting. Water Res. 46, 907–919. doi:http://dx.doi.org/10.1016/j.watres.2011.11.076

3. Delvigne, F., Goffin, P., 2014. Microbial heterogeneity affects bioprocess robustness: Dynamic single-cell analysis contributes to understanding of microbial populations. Biotechnol. J. 9, 61–72. doi:10.1002/biot.201300119

4. Delvigne, F., Zune, Q., Lara, A.R., Al-Soud, W., Sorensen, S.J., 2014. Metabolic variability in bioprocessing: implications of microbial phenotypic heterogeneity. Trends Biotechnol. 32, 608–616.

5. Gladden, J.M., Allgaier, M., Miller, C.S., Hazen, T.C., VanderGheynst, J.S., Hugenholtz, P., Simmons, B.A., Singer, S.W., 2011. Glycoside hydrolase activities of thermophilic bacterial consortia adapted to switchgrass. Appl Env. Microbiol 77, 5804–12. doi:10.1128/AEM.00032-11

6. Günther, S., Koch, C., Hübschmann, T., Röske, I., Müller, R.A., Bley, T., Harms, H., Müller, S., 2012. Correlation of Community Dynamics and Process Parameters As a Tool for the Prediction of the Stability of Wastewater Treatment. Environ. Sci. Technol. 46, 84–92. doi:10.1021/es2010682

7. Kinet, R., Dzaomuho, P., Baert, J., Taminiau, B., Daube, G., Nezer, C., Brostaux, Y., Nguyen, F., Dumont, G., Thonart, P., Delvigne, F., 2016. Flow cytometry community fingerprinting and amplicon sequencing for the assessment of landfill leachate cellulolytic bioaugmentation. Bioresour. Technol. 214, 450–459. doi:10.1016/j.biortech.2016.04.131

8. Koch, C., Günther, S., Desta, A.F., Hübschmann, T., Müller, S., 2013. Cytometric fingerprinting for analyzing microbial intracommunity structure variation and identifying subcommunity function. Nat. Protoc. 8, 190–202. doi:10.1038/nprot.2012.149

9. Koumoutsi, A., Chen, X.-H., Henne, A., Liesegang, H., Hitzeroth, G., Franke, P., Vater, J., Borriss, R., 2004. Structural and Functional Characterization of Gene Clusters Directing Nonribosomal Synthesis of Bioactive Cyclic Lipopeptides in Bacillus amyloliquefaciens Strain FZB42. J. Bacteriol. 186, 1084–1096. doi:10.1128/JB.186.4.1084-1096.2004

10. Mansfield, S.D., Mooney, C., Saddler, J.N., 1999. Substrate and enzyme characteristics that limit cellulose hydrolysis. Biotechnol Prog 15, 804–816.

11. Martin-Ryals, A., Schideman, L., Li, P., Wilkinson, H., Wagner, R., 2015. Improving anaerobic digestion of a cellulosic waste via routine bioaugmentation with cellulolytic microorganisms. Bioresour. Technol. 189, 62–70. doi:10.1016/j.biortech.2015.03.069

12. Martins, B.M., Locke, J.C., 2015. Microbial individuality: how single-cell heterogeneity enables population level strategies. Curr. Opin. Microbiol., Cell regulation 24, 104–112. doi:10.1016/j.mib.2015.01.003

13. Neumann, L., Scherer, P., 2011. Impact of bioaugmentation by compost on the performance and ecology of an anaerobic digester fed with energy crops. Bioresour. Technol. 102, 2931–2935. doi:10.1016/j.biortech.2010.11.068

14. Palmqvist, E., Hahn-Hägerdal, B., 2000. Fermentation of lignocellulosic hydrolysates. II: inhibitors and mechanisms of inhibition. Bioresour. Technol. 74, 25–33. doi:http://dx.doi.org/10.1016/S0960-8524(99)00161-3

15. Peng, X., Börner, R.A., Nges, I.A., Liu, J., 2014. Impact of bioaugmentation on biochemical methane potential for wheat straw with addition of Clostridium cellulolyticum. Bioresour. Technol. 152, 567–571. doi:10.1016/j.biortech.2013.11.067

16. Salehian, P., Karimi, K., Zilouei, H., Jeihanipour, A., 2013. Improvement of biogas production from pine wood by alkali pretreatment. Fuel 106, 484–489. doi:http://dx.doi.org/10.1016/j.fuel.2012.12.092

17. Scherer, P., Neumann, L., 2013. "Methano-compost", a booster and restoring agent for thermophilic anaerobic digestion of energy crops. Biomass Bioenergy 56, 471–478. doi:10.1016/j.biombioe.2013.05.021

18. Stewart, P.S., Franklin, M.J., 2008. Physiological heterogeneity in biofilms. Nat. Rev. Microbiol. 6, 199–210. doi:10.1038/nrmicro1838

19. Sutherland, I.W., 2001. The biofilm matrix – an immobilized but dynamic microbial environment. Trends Microbiol. 9, 222–227. doi:10.1016/S0966-842X(01)02012-1

20. Tachaapaikoon, C., Kosugi, A., Pason, P., Waeonukul, R., Ratanakhanokchai, K., Kyu, K.L., Arai, T., Murata, Y., Mori, Y., 2012. Isolation and characterization of a new cellulosome-producing Clostridium thermocellum strain. Biodegradation 23, 57–68. doi:10.1007/s10532-011-9486-9

21. Verstraete, W., Wittebolle, L., Heylen, K., Vanparys, B., de Vos, P., van de Wiele, T., Boon, N., 2007. Microbial Resource Management: The Road To Go for Environmental Biotechnology. Eng. Life Sci. 7, 117–126. doi:10.1002/elsc.200620176

22. Wittebolle, L., Marzorati, M., Clement, L., Balloi, A., Daffonchio, D., Heylen, K., De Vos, P., Verstraete, W., Boon, N., 2009. Initial community evenness favours functionality under selective stress. Nature 458, 623–6. doi:10.1038/nature07840

23. Wongwilaiwalin, S., Rattanachomsri, U., Laothanachareon, T., Eurwilaichitr, L., Igarashi, Y., Champreda, V., 2010. Analysis of a thermophilic lignocellulose degrading microbial consortium and multi-species lignocellulolytic enzyme system. Enzyme Microb. Technol. 47, 283–290. doi:10.1016/j.enzmictec.2010.07.013