

UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

**Sede Amministrativa: Università Degli Studi Di Padova**

Dipartimento Cribi - Centro Ricerche Biotecnologie Innovative

SCUOLA DI DOTTORATO DI RICERCA IN BIOCHIMICA E  
BIOTECNOLOGIE  
INDIRIZZO BIOTECNOLOGIE  
CICLO XXIV

***METAGENOMICS APPROACHES TO ANALYZE MICROBIAL BIODIVERSITY***

**development of new technologies  
and their subsequent application to environmental studies and bioreactor  
optimization**

**Direttore della Scuola** : Ch.mo Prof. Giuseppe Zanotti

**Coordinatore d'indirizzo**: Ch.mo Prof. Giorgio Valle

**Co-Supervisore** : Ch.mo Prof. Andrea Squartini

**Co-Supervisore** : Ch.mo Prof. Giorgio Valle

**Dottorando** : Riccardo Rosselli



# Tables of Contents

Abstract .....	7
Riassunto .....	8
Introduction .....	9
Chapter 1: Identification and classification of Bacteria.....	12
1.1 Criteria to identify and classify microorganisms .....	13
1.1.1 Identification of unculturable bacteria .....	15
1.1.2 16S rRNA gene analysis .....	16
1.1.3 Operational Taxonomic Units (OTUs).....	19
1.2 Principal features of the 16S ribosomal RNA.....	20
1.2.1 Characteristics of the ribosomal RNA .....	21
1.2.2 The ribosomal operon .....	23
1.2.3 The rRNA transcription process.....	23
1.2.4 rRNA genes evolution .....	26
1.2.5 Ribosomal RNA genes as markers for bacterial identification .....	27
1.2.6 Cons of a PCR-based approach.....	30
1.3 Development of universal primers and sequencing data analysis.....	30
Chapter 2. The Anammox bacteria .....	32
2.1 Nitrogen in nature .....	33
2.1.1 The nitrification process.....	34
2.1.2 The denitrification process.....	35
2.2 Biological approaches to nitrogen removal.....	36
2.3 The Anammox bacteria .....	39
2.4 A deep branch of prokaryotes: the Planctomycetes phylum .....	40
2.4.1 Anammox bacteria: morphology, physiology, genomics .....	41
2.4.2 Biochemistry of the Anammox reaction .....	45
2.4.3 Potential application of the Anammox process .....	46
2.4.4 Pros and cons of anammox reactions in industrial applications ..	48
2.5 Aim of the research .....	49
Chapter 3. The biological system of the dromedary rumen.....	50
3.1 The ruminal microbiota.....	52
3.2 Ruminal bacteria: Fibrobacter, Ruminococcus and Butyrivibrio.....	55
3.2.1 <i>Fibrobacter</i> sp.....	55
3.2.2 <i>Ruminococcus</i> sp.....	58
3.2.3 <i>Butyrivibrio</i> sp.....	60
3.3 The digestion mechanism.....	62
3.3.1 Cellulose structure and cellulolytic activity .....	62
3.3.2 Protein hydrolysis and digestion .....	66
3.4 The dromedary rumen .....	67
3.5 Aim of the research .....	69

Chapter 4.	Materials and Methods .....	71
4.1	List of Abbreviations and Terms Used .....	71
4.2	General Solutions and Preparations.....	72
4.3	Sample preparation for electron microscopy.....	73
4.3.1	Scanning Electron Microscopy.....	73
4.3.2	Transmission Electron Microscopy .....	73
4.3.3	Negative staining .....	74
4.4	Samples analysed.....	74
4.4.1	The dromedary rumen .....	74
4.4.2	The Anammox samples .....	74
4.4.3	Milan pilot bioreactor .....	74
4.5	Rumen and Anammox samples: total DNA extraction.....	76
4.5.1	Protocol Modifications .....	77
4.5.2	Dromedary rumen DNA extraction protocol.....	78
4.5.3	Anammox and Milan bioreactors DNA extraction.....	79
4.6	Polymerase Chain Reaction (PCR) .....	80
4.6.1	Semiquantitative PCR .....	82
4.6.2	PCR primers for ribosomal genes.....	82
4.6.3	Universal primers definition .....	83
4.6.4	PCR-products purification.....	85
4.6.5	PCR product lyophilization .....	86
4.7	RFLP.....	86
4.8	Bioinformatic tools .....	87
4.8.1	BLAST .....	87
4.8.2	ClustalW .....	88
4.8.3	CdHIT.....	88
4.8.4	Perl.....	89
4.8.5	MySQL .....	90
4.8.6	Online tools: Camera, Galaxy and RDP.....	90
Chapter 5.	16S rRNA genes analysis .....	91
5.1	Comparison between 16S rRNA genes from sequenced bacteria .....	91
5.2	Cluster analysis and sequences annotation.....	101
Chapter 6.	Anammox bioreactor: results and discussion .....	107
6.1	Electron microscopy .....	107
6.1.1	Scanning Electron Microscopy observation.....	108
6.1.2	Negative Staining Electron Microscopy.....	109
6.1.3	Transmission Electron Microscopy .....	110
6.2	Semiquantitative PCR reaction.....	111
6.2.1	Restriction Fragment Length Polymorphisms analysis.....	114
6.3	Sequencing .....	115
6.3.1	Sanger sequencing .....	115
6.3.2	Roche-454 sequencing.....	116

6.3.3	Cluster analysis .....	118
6.3.4	Rarefaction analysis .....	120
6.4	Sequences annotation and analysis of sample-related biodiversity ....	123
6.5	Genomic and biochemical data comparison .....	131
6.6	Virtual fingerprinting .....	133
6.7	Conclusions .....	136
Chapter 7.	The dromedary rumen associated microflora.....	138
7.1	Sequencing results and data analysis .....	138
7.1.1	Cluster analysis .....	139
7.2	Sequences annotation .....	142
7.3	Sequences annotation and analysis of sample-related biodiversity ....	146
7.4	Microbiological considerations.....	150
Bibliography.....		154



## **Abstract**

Metagenomics represents a recent approach to better analyze and characterize the environmental microbial biodiversity. The great advances in the field of nucleic acids sequencing allow to deeply analyze a microbial community focusing on specific genetic markers, the 16S rRNA genes.

In this work, the ribosomal genetic sequences of two bacterial population samples were deciphered using high throughput new generation sequencing.

A consortium of Annamox bacteria capable of metabolize nitrogen-based compounds was analyzed. These organisms are strongly interesting because of their potential application in bioremediation of nitrogen-based pollutants.

Several tools to follow the community development and to identify related species populating a well-defined Annamox system were developed.

Moreover, the dromedary rumen, *Camelus dromedarius*, was analyzed as example of natural bioreactor containing a microbial population specialized in the vegetal fibers transformation. From this study, several cellulolytic active bacteria that will become a potential resource for new biotransformation processes were found, suggesting a possible application in the field of livestock feed production.

## Riassunto

La metagenomica rappresenta un nuovo approccio finalizzato all'analisi e alla caratterizzazione della biodiversità microbica presente in un ambiente. Grazie allo sviluppo della tecnologie legate al sequenziamento, è possibile condurre un'analisi approfondita di un microbioma sfruttando marcatori genomici specifici come i geni 16S rRNA.

In questo elaborato, sono state analizzate le popolazioni batteriche di due campioni basandosi sulle sequenze di geni ribosomali ottenute tramite sequenziamento ad alta produttività'.

L'analisi di un bioreattore pilota in cui è stato selezionato un consorzio di batteri in grado di metabolizzare ammonio, ha permesso di identificare una popolazione Anammox. La potenzialità di questi organismi è rivolta allo sviluppo di nuove strategie per abbattere i livelli di composti azotati che si accumulano come scarto da diversi tipi di lavorazione industriale. Sono stati inoltre messi a punto strumenti validi sia per l'identificazione di nuove specie correlate che per il monitoraggio di un sistema nel quale l'equilibrio è già instaurato.

Inoltre è stato analizzato un bioreattore naturale rappresentato dalla cavità ruminale del dromedario, *Camelus dromedarius*, in cui una comunità si è specializzata nella trasformazione di substrati, quali le fibre vegetali. L'interesse nasce dalle numerose e potenziali applicazioni, tra le quali l'impiego nella produzione di mangimi per gli animali da allevamento. L'analisi ha infatti evidenziato la presenza di numerose specie note per la loro grande capacità' cellulolitica e quindi di una potenziale risorsa per nuovi processi di biotrasformazione.

## Introduction

Bacteria and Archaea are the most widespread life forms in the world: they can inhabit every environment of the planet, no matter how inhospitable it might be. They colonize places in which the conditions are prohibitive for most other organisms due to the ability in biotransformation processes deriving energy from a huge variety of substrates. The secret of the evolutionary success of bacteria is correlated to a simple cell organization, where a single chromosome occurs for the basal physiological functions and a variety of mobile elements allows a rapid response to the environmental variation.

Such a wealth of capabilities has prompted efforts to exploit microbial communities able to carry out specific and useful bio-conversions. A bioreactor is a system that supports a biologically active environment. It involves bacterial and/or eukaryotic cells, or biochemically active substances derived from such organisms and can perform the production, transformation or degradation of organic and inorganic compounds.

The detection and characterization of these communities is restricted by the difficulty to culture such organisms from complex samples (Gilbride et al., 2006) but the technological development partially solved these problems opening the way for new research.

Metagenomics and metatranscriptomics represent two recent fields of molecular biology which has enabled the examination of microbial diversity and the detection of specific organisms without the need of cultivation, by scanning their genetic material. The prefix meta is a Greek word that means 'transcendent'. This new science tries to explore the biological environment from a holistic point of view, transcending the individual life form, to focus on the genes that characterize the microbial community, carrying out the main reactions and functions in a particular environment. The entire biological systems could be viewed as a single super-organism in which several interacting communities modify the environment and themselves.

The advancement of genomics and the analysis of genomes by sequencing and

annotation of genes, associated with the recent advances in bioinformatics (softwares and databases), have facilitated the understanding of biodiversity and of the potential functions of bacterial consortia (Mavromatis et al., 2007). Thanks to the development of Next Generation Sequencers it is now possible to obtain a huge number of sequences directly from DNA, PCR products or retrotranscribed-RNA (Morozova et al. 2008). These can partially solve classical sequencing-associated problems, such as clone library construction, and, due to the considerable high-throughput, can reduce time and money otherwise required by the standard Sanger DNA sequencing technology (Hall, 2007).

Other DNA-based techniques such as amplified ribosomal restriction analysis (ARDRA), ribosomal spacers analysis (RISA), terminal restriction fragment length polymorphisms (t-RFLP) and denaturing gel electrophoresis, are used in many fields of community microbiology as flanking tools or preludes to the sequencing techniques (Talbot et al., 2008).

Each information concerning the environment and its living organisms should be analyzed and compared, to find a correlation between causes and effects and to understand which mechanisms sustain the biological system. Understanding which mechanisms regulate the structure and function of a microbial community can be very useful for the application of such consortia in man-driven processes.

As ideal bioreactor systems we envisage two examples:

- An industrial pilot plant hosting a consortium of undetermined bacteria originating from nitrification-denitrification sludges and evolved under inorganic feeding conditions. Such incubation is conducive to the development of an anammox-type consortium of biota able to convert ammonia into nitrite and further to gaseous dinitrogen. Such application is of primary interest in the development of novel strategies to achieve the abatement of nitrogen-containing industrial wastes. A characterization of these bacteria would also foster the setup of patents for the benefit of regional applications independent of the

present expensive alternatives based on unculturable bacteria marketed from private foreign companies.

- As a second type of bioreactor we planned to explore a more complex community isolated from a natural habitat, the camel rumen, which is an extremely efficient biosystem of organic matter transformation, prone to reveal an array of metabolic diversity with great potential for industrial applications. A collaboration with the University of Constantine, Algeria has allowed the study of these animals. Camels live in semi wild conditions and feed on xerophytic desert weeds with very high cellulose and lignin content. Both the naturally occurring microbiota of the camel rumen and the ones evolving in subsequent in-vitro enrichment in bioreactors containing agro industrial residues was considered for the analyses.

## **Chapter 1: Identification and classification of Bacteria**

Microbial classification, is a necessity for scientists and workers dealing with microorganisms. The goal, is to provide a practical framework in which the relationships among organisms follows a logical and hierarchical structure and defined groups display a similar pattern giving an easier overview of the biological complexity.

Because of the medical relevance of pathogenic microorganisms, bacterial taxonomy has often followed preferential paths and the identification of an organism has been considered in light of a possible therapeutic concept, bypassing the formalisms concerning other relational aspects between species. This consideration remarks the double-faced aspect of this field of research. Identification and classification represent the dichotomy of bacterial taxonomy, where the first branch provides a pragmatic and functional grouping of species and the second one takes into account the hypothetical evolutionary relationships between organisms to group them.

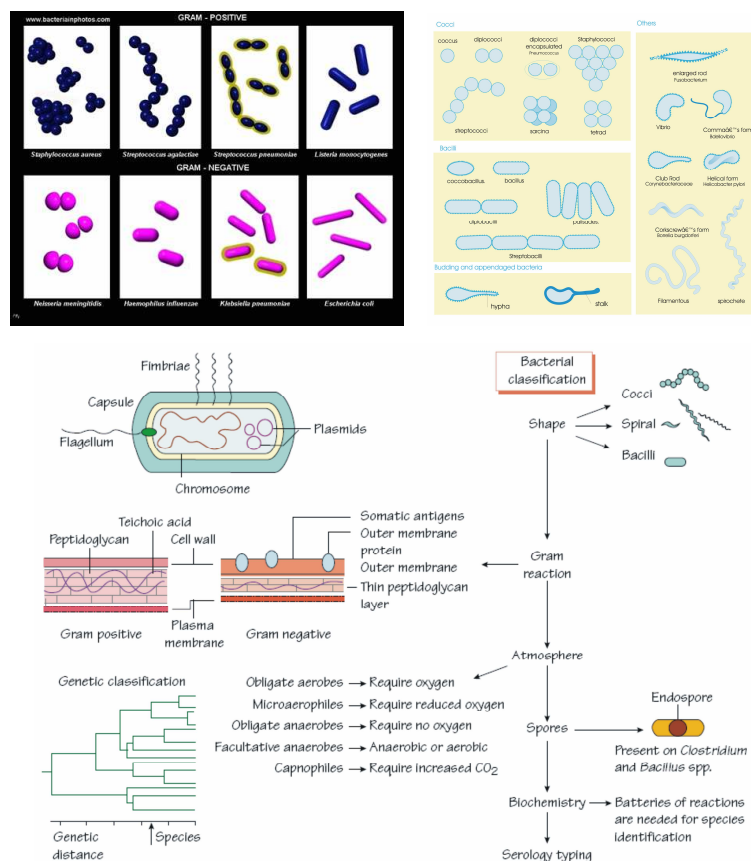
Presently, new perspectives concerning the use of microbial cells industrial applications, bioreactors for several man-driven processes, bioremediation, have led a re-evaluation of the bacterial classification principles. This requirement, is determined by the need of the highest possible level of information concerning the identification and characterization of the bacterial communities that are operative in such applications.

Bacterial classification is closely intertwined with the development of new technologies and faces difficult tasks due to the complexity of microbiological studies, the impractical dimension of organisms, the need of cultivation and isolation to obtain a useful and analyzable sample.

## 1.1 Criteria to identify and classify microorganisms

Left aside the historical views of the microbial world that for almost three centuries has fascinated the scientific community, the first rigorous descriptions date to the late nineteenth century, where a morphology-based analysis led to the description of the first genera of bacteria as members of the plants. For several years morphological classification represented the main strategy to give a description of microorganisms.

In 1925, the definition of prokariotes and eucaryotes appeared (“*Pansporella perplexa: amoebian à spores protégés parasites des daphnies: réflexions sur la biologie et la phylogénie des protozoaires*”, Chatton, 1925) and in 1961, when some groups had already been defined (including Actinomycetes, Myxobacteria, Spirochaetes), the definition of *Bacterium* was formalized producing the definitive subdivision of these groups of organisms.



**Fig.1.** Features used for bacterial classification: Gram staining, morphological information and biochemical data. (sources: www.bacteriainphotos.com, www.blackwellpublishing.com).

Chemotaxonomy, based on the specific chemical composition of several cellular components that characterize different groups of organisms, as the chemical composition of the cell wall of Gram positive and Gram negative bacteria, provides other parameters for bacterial classification. A chemical profile suited to emphasize several enzymatic activities (as a measurement of color or phase changes in a culture medium due to an indicator molecule or medium composition) can lead to an identification based on specific metabolic pathways (Fig. 1)(Elston et al., 1971; Devenish et al., 1980).

On one hand, the sequencing of an increasing number of bacterial genomes and the development of bioinformatic tools for the analysis, represent primary data sources for prokaryotic classification and have allowed to partially simplify the 'hard marriage' between identification and classification. On the other hand, it is known that the natural development of new bacterial species is determined by vertical inheritance and, perhaps even more, affected by lateral evolution and horizontal gene transfer (Ochman et al., 2000). Genomic discoveries opened the way for new evolutionary scenarios but the latest definition of the prokaryotic phylogeny represented by a network and not by a tree (Kunin et al., 2005), increases the difficulty in finding a link between classification, identification and definition of historical relationships among species.

A useful suggestion is derived from the increasing number of more accurate informations, molecular data and comparative analysis. Nowadays, a scientific recognized classification model is based on several informations related to different microbial features and defined as *polyphasic approach* (Colwell, 1970)

To achieve a sound classification, phenotypic, genotypic and phylogenetic evaluations (often based on sequence comparison) must be considered to yield robustly supported taxonomical identification of specific organisms.

- **Phenotypic features:** These comprise morphology-related cellular characteristics as shape, Gram staining, structures to facilitate motility (pili and flagella) and their position, as well as colony characteristics as dimension, form, surface and consistency. Phenotypic features include cell wall-related molecules and specific molecular patterns related to fatty acids,

quinones or other substituents. Biochemical features involve the capability to grow in specific conditions of temperature, pH, salt concentration or evidence of precise metabolic pathways.

- **Genotypic features:** this approach includes all the DNA- or RNA- directed analysis. These methods dominate the modern microbiological studies due to the development of next generation sequencers and the new opportunities supplied by the nucleic acids direct sequencing (DNA, RNA or PCR products). Considering the genome composition, a parameter included in several description of bacterial clades is represented by the DNA base ratio taking into account the G+C percentage on the whole genomes. This measure is still useful for the bacterial identification at higher level (e.g. the phylum). The percent DNA-DNA hybridization and its thermal stability may be used for bacterial classification and species definition. Measuring DNA-DNA hybridization value and percentage binding, an indication concerning sequences similarity can be obtained if the relation for which each mismatch can determines a decreasing in thermal stability (from 1% to 2%) is considered. Currently, DNA-DNA hybridization methods have been replaced by 16S rRNA sequencing and comparison. Sequence similarity used for DNA-DNA hybridization analysis remains the reference parameter for 16S rRNA comparison and “70% or greater DNA-DNA relatedness” (Wayne et al. 1987) which defined species, corresponds to the 97% of similarity required by a 16S rRNA comparison for the same taxonomic level definition.

### **1.1.1 Identification of unculturable bacteria**

Application of all the available criteria, a polyphasic approach, represent the most accurate way for bacterial identification and clasification (Okabe et al., 2010) but the scientific world knows that a huge number of microorganisms representing the majority of biodiversity is recalcitrant to culturing and isolation (Amman et al, 1995)

It is estimated that only 1% of bacteria can be isolated by cultivation (Pace, 1997; Torsvik et al., 2002; Vartoukian et. al, 2010). The number of cells which can be

observed using microscopy, is often two orders of magnitude higher with respect to the number of colonies obtained inoculating the same sample in culture enrichment media. This simple observation, defined as “great plate count anomaly” (Staley et al., 1985), is the result of the growth recalcitrance due to several occurrences. For growth, an organism may depend on the metabolic activities of other free species (Qiu et al., 2003) or requires a specific association in a biofilm-like structure (Stoodley et al., 2002). It may be characterized by slow growth rate or needs specific conditions in terms of pH, temperature, salinity (Kopke et al., 2004). Within 52 currently proposed major lineages (at phylum level), at least 26 are characterized by unculturable organisms (Hugenholtz, 2002; Rappé et al., 2003).

All these features, underline a limit in the true estimation and understanding of microbial biodiversity. Alternative approaches that bypass the need of bacterial cultivation and isolation must be applied to identify and classify microorganisms and characterize their populations.

### **1.1.2 16S rRNA gene analysis**

Nowadays, a fundamental step for bacterial classification is represented by sequencing or fingerprinting-based protocols on whole genomes, and on 16S rRNA genes as principal and easily detectable marker for bacterial identification. Other housekeeping genes have the fundamental requirement and evolutionary characteristics similar to 16S rRNA genes. The RNA polymerase B-subunit gene (*rpoB*) can represent an alternative to ribosomal genes as molecular marker (Walsh et al., 2004; Case et al., 2007) while other genes can be used as markers for specific taxa (Achenbach et al., 2001).

The study of these ubiquitous and widely distributed markers on nucleic acids directly extracted by the sample without the need of bacterial cultivation, has revolutionized microbiology and the knowledge of bacterial communities composition for several environments.

In particular, analysis from directly extracted 16S rRNA (Lane et al., 1985), PCR development and application to microbial population studies (Giovannoni et al.

1990) fostered the discovery of unknown species in several different environments partially solving most of the prior problems related to the amount of starting material.

A large number of techniques involves PCR and requires the definition of a pair of primers for specific gene amplification. PCR products, can be employed for many experimental protocols in which the main purpose is to define a pattern, a sequence or any other parameter for the identification, clustering, comparison of the bacterial population members

- Pulsed-Field Gel Electrophoresis (PFGE): PFGE represents a method for separation of large DNA molecules due to a capacity of resolution up to 10 Mb.

The direction of electrical field changes periodically determining a different direction for DNA migration, thus solving the random coiling of large size DNA (such as entire chromosomes) and allowing the size-dependent distribution of molecules through the gel matrix yield a genome- or chromosome-dependent specific pattern (Schwartz et al. 1984; Herschleb et al., 2007).

- PFGE + ENDONUCLEASE DIGESTION: this technique is PFGE-based but cell lysis is followed by a digestion due to one or more rare-cutter restriction enzymes. Electrophoretic fragments distribution, reflects the presence of cutting sites in DNA allowing the definition of a specific pattern related to an organisms. PFGE on digested DNA is useful for the identifications of different strains of the same species (Tenover et al., 1995; Gordillo et al., 1993)
- Restriction Fragment Length Polymorphisms (RFLP): like all restriction-based approaches, the principle of RFLP techniques is the different length and distribution of DNA through a gel matrix related to the availability of specific restriction sites on the sequence.

Using PFGE-based analyses, a large amount of starting genomic DNA is required, making them useful for single species analysis and culturable bacteria. RFLP target sequences can be represented by PCR products,

solving problems related to the template amount and allowing the use of sequences obtained by universal primers directly on heterogeneous microbial populations. Due to this, RFLP is useful for the identification of microorganisms from strain to genus level.

- Denaturing Gel Electrophoresis (DGE): In DGE, amplified DNA products (commonly 16S rRNA genes) are separated along a polyacrylamide gel under a gradient of denaturing conditions determined by chemicals agents (Denaturing Gradient Gel Electrophoresis, DGGE), temperature (Temperature Gradient Gel Electrophoresis, TGGE) or temporal temperature gradient (Temporal Temperature Gradient Electrophoresis, TTGE). Denaturing conditions, determine a migration of DNA related to the different G+C content, distribution and melting properties (Myers et al., 1985; Fromin et al. 2002; Muyzer et al., 1993).
- Amplified Ribosomal DNA Restriction Analysis (ARDRA): ARDRA implies a ribosomal gene amplification followed by enzymatic digestion. This technique can be considered a variation of RFLP specific for rRNA genes and is very useful for microbial communities identification and monitoring (Martinez-Murcia et al., 1995; Smit et al. 1997).
- Sequencing: obtaining informations from DNA sequences or RNA represents the last frontier of microbiology. The amount of available genomes and sequences has undergone a remarkable surge during the last 20 years due to the development of sequencing technologies.

Until recently, using Sanger sequencing several prokaryotic genomes, PCR-products, environmental and cloned sequences have been analyzed. The development of new technologies applied in this field of research, has increased the reaction throughput and the ease of obtaining a huge amount of bases directly from nucleic acid extracted and manipulated from a sample. These breakthroughs, have partially solved and overcome some bottlenecks related to Sanger sequencing as the difficulty to clone some genes (because of their toxicity, Sorek et al. 2007) and the limit for which one sequence is obtained by one clone and creation of a clone library can be expansive in terms of amount of starting material, time and money.

### **1.1.3 Operational Taxonomic Units (OTUs)**

As described above, the aim of classification is to identify organisms grouping them in an historical picture represented by a tree or a more realistic network, hardly accepted by taxonomists.

The development of new technologies, the capability to obtain information from non culturable organisms as short fragments of their genomes or genes using fingerprinting protocols or sequencing and the difficulty in following a polyphasic approach, result in several questions related to the appropriate consideration and integration of such fragmented data.

Several online tools have been developed for the annotation of partial and complete sequences. Databases containing a huge amount of sequences are available and useful for the comparative study and annotation.

Another approach is being developed for the analysis, clarifying the interpretation of data relative to microbial population studies and allowing the comprehension of very complex ecosystems. In this approach data, sequences or restriction patterns (Liu et al., 1997; Dunbar et al. 1999) are assigned to Operational Taxonomic Units (OTUs), considering and comparing the single information within its specific data set (Schloss et al., 2011). An OTU-based analysis really represents a cluster-based approach without speculation on classification or identification results and taking in consideration the sample itself and not the comparison between data and other databases (Gevers et al., 2005).

Fundamental for this analysis is the data clustering step. For this purpose, several algorithms are available (e.g.: CD-HIT, [www.bioinformatics.org/cd-hit](http://www.bioinformatics.org/cd-hit)).

Once OTUs identity and number are defined, several statistical approaches can be applied for the data analysis. Rarefaction and richness estimators, can reveal information on community but also on the investigation method efficiency suggesting if the data coverage is sufficient to describe the entire biodiversity. Nonparametric estimators stemming originally from the mark-release-recapture (MRR) statistics, are preferred with respect to the parametric estimators because they do not require abundance relative data and can be applied at small dataset (Hughes et al., 2001; Curtis et al., 2002).

- Rarefaction curve: a rarefaction curve is a representation of the fraction of OTU obtained by an experimental procedure with respect to the total number of OTU potentially contained in the entire sample. This evaluation takes into account common and rare representatives (Koellner et al., 2004; Wooley et al., 2006) yielding a measurement about the knowledge depth achieved on a specific population (Kenneth et al., 1975).
- Shannon index: widely used, relates to species richness defining a different weight for common or rare species (Hill et al., 2003). A limit inherent to this parameter, is the requirement of a clear identification and classification of each single species (category) (Torsvik et al., 1990) which is one of the most difficult microbiological goals as described above.
- Abundance Coverage Estimator (ACE): this nonparametric index considers the number of rare OTUs (less than 10) and the number of single OTUs to evaluate how many more undiscovered OTUs could be present in the sample (Chao A., 1987). This index, can give an estimation about the number of sequences (for example) that should be obtained to cover the entire biodiversity of an environment (Chao A., 1992).
- CHAO 1 index: with the Chao 1 nonparametric estimator, derived from mark-release-recapture indexes (Chao A., 1984), it is possible to infer information on the total species richness considering singleton and doubleton OTUs (appearing once or twice respectively in the sample).

## **1.2 Principal features of the 16S ribosomal RNA**

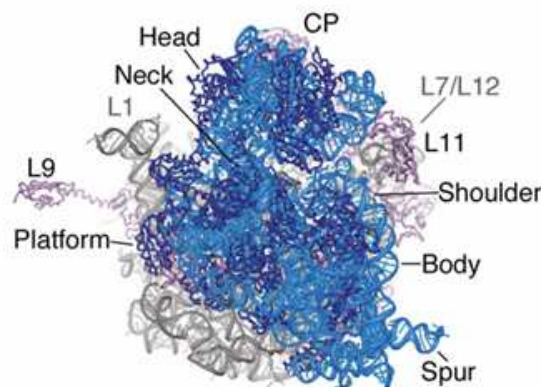
As recognized by Woese during the 70's and developed by several authors to date, rRNA analysis by sequencing or fingerprinting, represents a fundamental step for bacteria, archaea and eukarya identification and an essential issue for classification (Woese et al., 1975; Woese, 1987).

In all living organisms, protein biosynthesis at the basis of cellular growth and development is catalyzed by the ribosomes making these one of the most widely distributed structures.

### 1.2.1 Characteristics of the ribosomal RNA

Functional bacterial ribosomes are constituted by a small subunit, 30S, and a large subunit, 50S. The 30S subunit is the result of the close interaction between the 16S rRNA, 1500 bases long, and 21 proteins folded in a ~900 kDA structure.

The 50S subunit consists of 23S rRNA, 2300 bases long organized in a more compact structure than 16S rRNA (Lee et al., 2007). In addition, the 5S rRNA is 150 bases long (Szymanski et al., 2003). Nucleic acid chains of the 50S subunit, interact with 31 proteins (Woodson et al., 1998). For all the rRNA chains, sequences involved in the protein interaction are named H and ribosomal associated proteins are identified by S letter followed by an ascendent number (Schuwirth, 2005).



**Fig. 2.** *E. coli* 70S ribosome. 30S subunit and associated proteins are colored in blue while 23S and 5S RNAs and proteins are colored gray and magenta. (Schuwirth, 2005).

A series of twelve connecting bridges, maintain the assembly between the two subunits. Six of them are highly conserved among all the domains: one is determined by a protein-protein interaction whereas the others are determined by RNA-RNA or RNA-proteins interactions (Liiv et al., 2006).

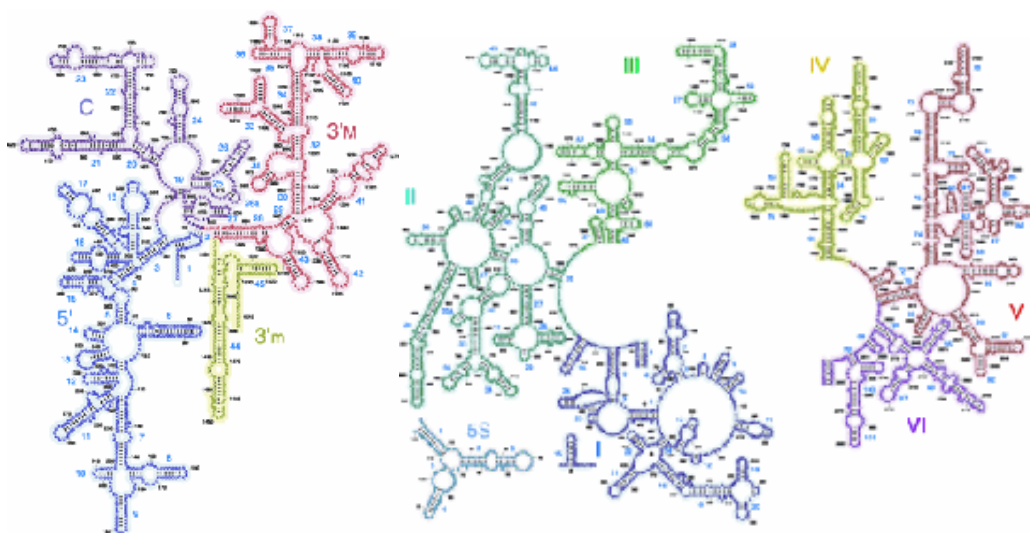
The 30S ribosomal subunit plays an essential role during mRNA translation, monitoring correct pairing between messenger and the appropriate tRNA. 50S subunit is involved in the peptide bond (Wimberly et al., 2000),

Because of the fundamental role of ribosomes in ensuring cell survival, a high selective pressure is focused on the key sites fundamentals for their function,

determining a specific evolution which is characteristic for their sequences. Infact, rRNA and proteins interaction are possible due to the specific secondary structure, stem, loop and bulged, which is a consequence of nucleic acid chains folding (Smith et al., 2008).

Specific ribosomal chains residues are essential to determine the correct tRNA selection and proteins synthesis accuracy (McClory et al., 2010). The nucleic acid sequence represents a fundamental trait in determining the differential stiffness (Gao et al., 2003) at the basis of ribosomal conformational changes, modulating the decoding process (Rodnina et al., 2002) and conferring a direct role in translational accuracy and fairness (Arkov et al., 1998, Noller et al., 2006).

Ribosomal RNA folding is the consequence of bases association, protein interaction and thermodynamic parameters which are still under investigation (Gutell et al., 2002) . These complex interactions involved common Watson-Crick pairing but also other interactions (Wayne et al., 2002), like Hoogs-teen edge and Sugar edge that link other residues such the 2'-hydroxyl group of sugars (Leontis et al., 2001). Specific conserved nucleotides are modified after transcription: isomerization of uridine to pseudouridine (Baxter-Roshek et al., 2007) and methylation characterize several residues in the most functionally relevant regions (Bakin et al., 1993, Yusupof et al., 2001, Chow et al., 2007).



**Fig. 3.** Secondary structure of 16S, 23S and 5S rRNAs of *T. thermophilus*. 16S: In blue and magenta, red and yellow are underlined the 5', central, 3'-major, 3'-minor domains respectively. (Yusupov, 2001).

### **1.2.2 The ribosomal operon**

Nucleic acid components, just represented by the three rRNA chains, are transcribed starting from a structurally highly conserved operon.

From 1 to 16 rRNA operons are interspersed in the bacterial genome, probably reflecting the different adaptive strategies between bacterial species (Shrestha et al., 2007). High copy numbers of these operons, are found in species characterized by a rapid response to variable growth conditions (Klappenbach et al., 2000) whereas a low copy number is characteristic of organisms that inhabit stable environments, as for example, several obligate pathogens (Stevenson et al., 2004, Lee et al., 2008).

Ribosomal operons are often located in the proximity of the origin of replication and transcribed in the same direction of DNA synthesis. This has a double functional meaning: firstly, it allows a continuous transcription process in balance with the leading DNA strand synthesis. Secondly, ensures a sufficient amount of functional ribosomes during the cell division, when the replication process runs faster than the cell division and bacterial cell must support a transient poliploidy condition (Nierhaus et al., 2004).

The relevance of rRNA for cell survival, can be also evaluated considering the presence of fragmented rRNA in the bacterial cell. 16S and 23S have specific cleavage sites that determine the presence of uncompleted forms of functional rRNA. Cleavage occurs in several regions, often located at the 5' end of 16S rRNA or other taxa-specific sites of 23S. Despite the uncompleted sequences, the transcription complex maintains the full functionality underlining the capability of rRNA sequence to determine successful of protein synthesis (Evguenieva-Hackenberg, 2005).

### **1.2.3 The rRNA transcription process**

Ribosomal RNA transcription is a well recognized process for several bacterial taxa. It is remarkable that, despite the similarity between sequences derived by the common evolution-determined pattern between distantly related organisms and the common series of reactions, the enzymes involved in several step of this

process are different.

All bacterial rRNAs are transcribed starting from several promoters under a common regulatory mechanism and as a unique polycistronic pre-rRNA (Murray et al., 2003; Shneider et al., 2003). Starting transcription from each promoter appears related to the cell growth rate (Josaitis et al., 1995).

The rising molecule is processed due to the action of several nucleases and assembled in a mature rRNA with the ribosomal proteins.

At the basis of the process, the 30S rRNA synthesis is characterized by a structural isomerization that determines the formation of pseudoknots, the secondary structure core of 16S rRNA (Brink et al., 1993). RNase III is the first highly conserved enzyme involved in the cleavage process: this enzyme recognizes specific double stranded sites determined by a base-pairing between 16S and 23S (Young et al., 1978). Such activity, determines the three ribosomal precursors release. Also any tRNA transcribed starting from the hypervariable 16S-23S spacers are processed during this reaction (Gurtler et al., 1996). In *Escherichia coli*, where the mechanism of maturation has been well investigated, a precursor of 16S rRNA, 17S, is characterized by extra 5' and 3' nucleotide tails (Young et al., 1978). Cleavage of 5' extra nucleotides is catalyzed in a two step mechanism by the subsequent activity of the structurally related enzymes RNase E and RNase G which determine the formation of a 16.3S precursor of 16S (Wachi et al., 1999, Li et al., 1999, Ow et al., 2003). Finally, also 3' extra nucleotides are removed but the endonuclease involved in such specific step remain unidentified (Deutscher et al., 2006, Gutgsell et al., 2010).

Both extremities of 23S precursors are characterized by extra residues tails but the cleavage mechanism is still under investigation. RNase G seems to be involved in the 5' end cleavage (Song et al., 2011) whereas the exoribonuclease RNase T, a nuclease related to the proofreading domains of bacterial DNA polymerase III (Koonin et al., 1993), removes the 3' extension (Li et al., 1999).

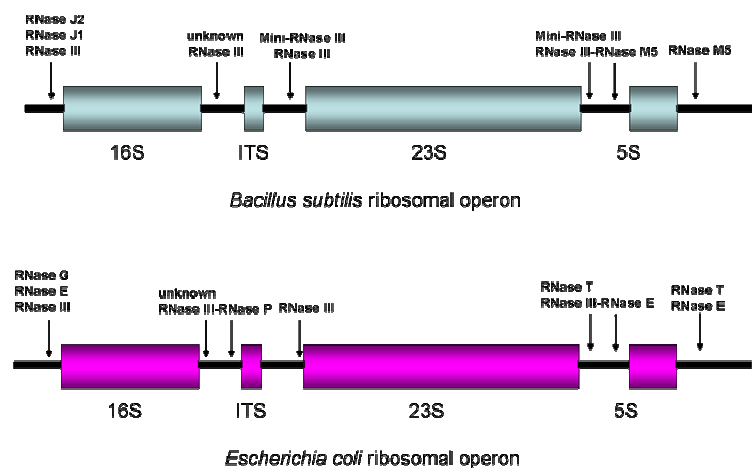
The last 5S pre-rRNA and eventually co-transcribed tRNA at the ribosomal 3' end (Gegenheimer et al., 1977) are separated by the 23S rRNA due to the activity of RNase III as 9S rRNA. Such molecule is subjected to RNase E (Roy et al., 1983) that determines a partial reduction of the extra nucleotides at each end. Finally, the

RNase T, determines the cleavage and maturation at 3' end (Li et al., 1995) whereas the enzyme involved in the 5' maturation remain unknown.

While the cleavage of the entire pre-rRNA is similar to that in *Escherichia coli* in *Bacillus subtilis* other enzymes are involved (Herskovitz et al., 2000). Genomic studies on several Gram-positive bacteria, revealed the lack of any homologues to RNase E. Two enzymes have been identified with a similar effect to RNase E, RNase J1 and RNase J2 (Even et al., 2005). In particular, RNase J1 seems to plays the same role in the pre-rRNA maturation, determining the cleavage of the 5' extra nucleotides end of 16S rRNA (Britton et al., 2007; Mathy et al., 2007).

Also pre 23S rRNA maturation is different if compared with the process occurring in *Escherichia coli*. Both extra residues at 5' and 3' end of 23S are the target of another class of RNase III discovered in *Bacillus subtilis* but also found in other Firmicutes and Cyanobacteria named mini-RNase III or Mini-III for short (Redko et al., 2008, Olmedo et al., 2008).

In Firmicutes, maturation of the smallest 5S precursor of rRNA requires the Toprim domain related protein (Allemand et al., 2005) RNase M5 (Sogin et al., 1974; Condon, 2007). Such enzyme, is responsible of the cleavage that determines the formation of a functional 5S RNA in a single step due to the interaction with some ribosomal related proteins (Stahl et al., 1984).



**Fig. 4.** Comparison between RNase involved in the ribosomal maturation in the Gram negative *E. coli* and the Gram positive *B. subtilis*

#### 1.2.4 rRNA genes evolution

The strong selective pressure that acts on the rRNA genes and on their transcribed molecule, determines a mutation rate which is unique if considering nucleotides position and the implications for the secondary structure. Stems, loops, bulges and the junctions between them, are subjected to a particular evolutionary rate (Smit et al., 2007), related to the different effects that the variation may have on the ribosomes functionality.

Stems are structural components of rRNA that are poorly involved in direct interactions within the entire translational complex. The variability at this level is higher with respect to the unpaired regions represented by loops (Rzhetsky et al., 1995; Otsuka et al., 1999): this could suggest that every mutation able to maintain the base pairing in these regions may be possible. But a relation between G-C content of stems and organisms lifestyle (e.g. optimal growth temperature) indicates a base-related selection which was instead denied for the entire genomes limiting the mutation rate (Wang et al. 2002)

The evolutionary rate related to base variations, seems to be lower in the center of ribosomes which represents the catalytic core of proteins translation (Wuyts et al., 2001), at the level of specific modified nucleotides involved in rRNA folding (Helm et al., 2006) and structure stabilization (Meroueh et al., 2000). Highly conserved bases characterize single strand regions of the secondary structure responsible for example of the interaction between 30S and 50S subunits (Pulk et al., 2006). Moreover, regions which can directly determine an initiation of translation by the correct start codon, involved during the elongation processes of proteins synthesis (Sun et al., 2010), in monitoring codon-anticodon interaction and translational fidelity maintaining (Ogle et al., 2005) and right tRNA selection (McClory et al., 2010) are characterized by a very slow evolutionary rate.

Due to such occurrence, rRNA genes sequence is characterized by highly conserved regions which are similar for distantly related organisms and fundamental for the rRNA folding and protein interactions, and hypervariable regions which are species-specific but also strain or population specific.

### **1.2.5 Ribosomal RNA genes as markers for bacterial identification**

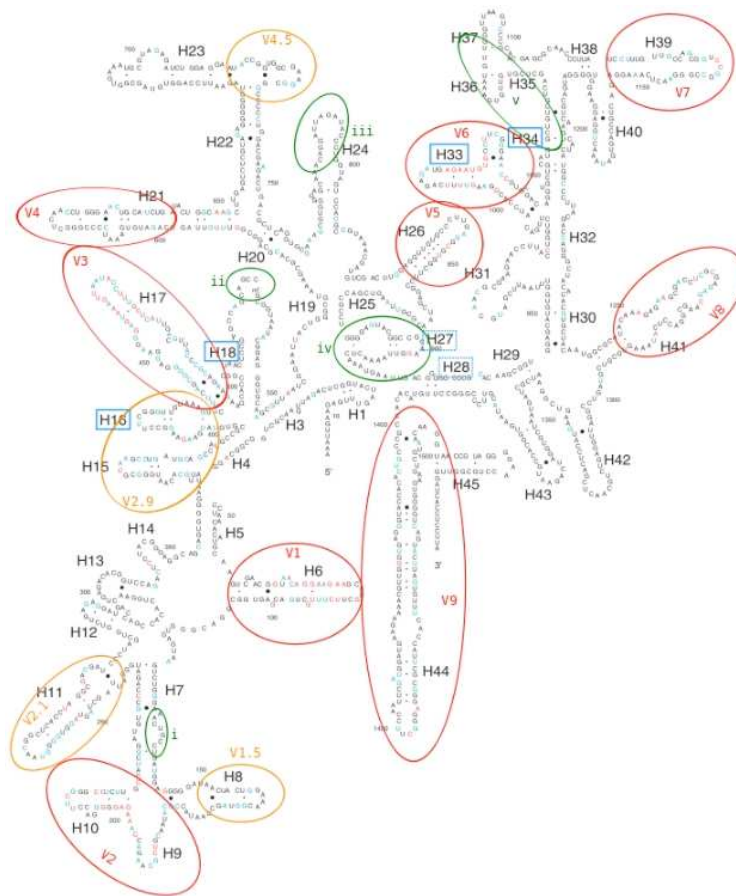
Both 5S, 16S and 23S have been used as markers for bacterial identification and classification.

The proof of principle of any rRNA analysis, represented by the wide distribution of such genes and the characteristic evolutionary rate, is known since the 70's (Woese et al., 1975).

Due to the smaller size and easier determination of primary structure, 5S rRNA was considered firstly as the main target for these analysis (Browlee et al., 1967; Sogin et al., 1971). Also 23S genes represent a good marker for bacterial classification. Greater length and more detailed information which can be obtained, make it the potentially better target for classification and phylogenetic studies (Hunt et al., 2006).

The development of sequencing technologies, the length compatibility with classical Sanger sequencing (few sequences needed to obtain the entire 16S rRNA gene assembly) that for more than 20 years represented the main strategy for genes and genomes sequencing, has rendered the 16S rRNA gene the 'golden standard' for this purpose, facilitating analyses, increasing accuracy and leading to the discovery of several new genera and species

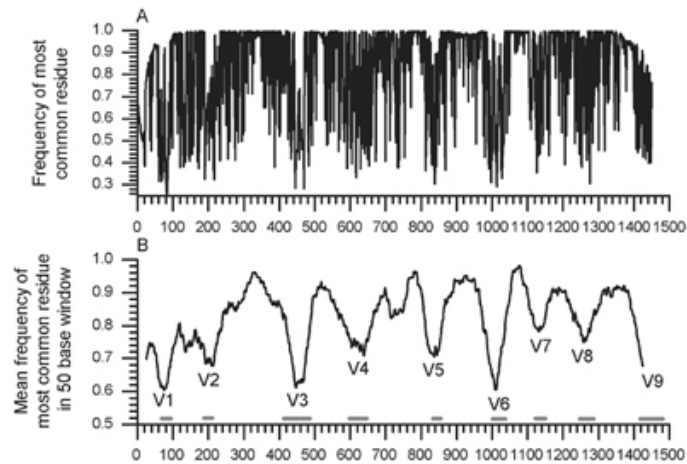
An estimation of substitution rate that characterized every region, reveals that in 16S rRNA genes the rate is ~7000 times higher for the variable regions with respect to the conserved regions (Van de Peer et al., 1996; Hashelford et al, 2005)



**Fig. 5.** Secondary structure of a prokaryotic 16S rRNA gene. Colour legend: red = hypervariable regions, yellow = variable but anonymous regions, green = conserved regions. (Adapted from Case, 2007).

Because of the great number of characters which are potentially under a strong selective pressure, it is anyhow rather difficult to define a unique evolutionary pattern for the entire set of rRNA genes.

For several bacterial 16S rRNA genes, nine hypervariable regions and eight high conserved regions were found and characterized, underlining the possibility of utilizing them for the identification of microorganisms (Chakravorty et al., 2007).



**Fig. 6.** Hypervariable region in the 16S rRNA gene. A: the frequency occurrence of the most common nucleotides at each base position. These frequencies represent the variability within the gene. B: The location of hypervariable regions. (Adapted from Hashelford, 2005).

Interspersed high conserved regions, determine the possibility of defining several pairs of universal primers able to give a wide length range of amplification products involving one or more hypervariable regions (Fig. 6).

200 bases (Wilck et al., 2001), 400 bases (Bosshard et al., 2003) the entire genes (Sacchi et al., 2002) have been used and it is remarkable that also 100 bases, if within the appropriate 16S rRNA hypervariable region, can give the same information than a full length template (Liu et al., 2007).

The efficiency of a single or groups of regions considered in providing informations about bacterial classification has been widely reported, for sequencing and for other 16S rRNA genes analysis as RFLP or DGGE.

V1-V2 (Balcazar et al., 2007), V1-V3, V3-V5, V4-V5 (Schwieger et al., 1998) V6-V8 (Nubel et al., 1996; Felske et al., 1996), V7-V8 (Ferris et al., 1996)

Phylogenetic analysis based on whole genome comparison with ribosomal genes are rather similar (Bansal et al., 2002), confirming the high potential of this application and the relevance of the primer definition process during the set up phases of the experiment.

As gold standard for bacterial identification, several molecular approaches can be applied for 16S rRNA genes analysis and bacterial classification, based on a

widely used laboratory technology, the Polymerase Chain Reaction.

### **1.2.6 Cons of a PCR-based approach**

On one hand a PCR-based approach represents a rapid and low-cost strategy to obtain informations about a microbial population on the other hand several bias can be introduced by these techniques.

Problems concerning sample handling, DNA extraction or primer specificity may determine an altered understanding of a biological system.

- Sample related bias: different cellular layer compositions may produce a different response to lysis methods and lysis recalcitrance may determines an under-rappresentation of some species. DNA quality and quantity can be related to extraction protocols (de Liphay et al., 2004) and PCR and enzymatic reactions may be altered by sample-related inhibitors o (Takhuria et al., 2008; Feinstein et al, 2009). Several analysis compared different extraction protocols evaluating nucleic acids yield and purity but a standard a strategy able to satisfy any type of experiment remain unknown.
- PCR related bias: Primers are defined on the basis of several 'a priori consumptions' represented by a sequences dataset. Only sequences considered in such database may be amplified by PCR and a huge amount of unknown informations remain undiscovered. Because of this problem, all PCR-based protocols, RFLP, DGE, ARDRA and sequencing bear an inherent problem that can not be solved.

### **1.3 Development of universal primers and sequencing data analysis**

The simplest but most effective strategy for bacterial identification and classification is represented by the definition of a pair of primers which can recognize the same conserved regions and amplify the hypervariable regions which will be used for the identification. Moreover, the singular evolutionary rate among bacterial lineages, allows the definition of primers for several bacterial

taxa, determining a specific clade-based analysis.

Next-generation sequencing represents the most powerful tool in providing information about a microbial community. Because of the huge amount of data, it must be necessarily developed together with appropriate bioinformatics tools which must provide an easy analysis to give a rapid and exhaustive interpretation of the results.

New discoveries about environments were determined by the progress of recent technologies but a problem is now represented by the huge amount of not annotated or poorly analyzed data.

The number of sequences of the Ribosomal Database Project increased in few years from 150000 to more than 1000000: less than 8000 queries can be downloaded if quality and annotation thresholds are used. Moreover, a huge number of not-identified bacteria may represent the output file if an online BLAST research at the NCBI site was carried out using a 16S rRNA sequence as query.

To allow identification and classification of a query, the definition of a similarity threshold represents a fundamental step: 97.5% of similarity sets the commonly accepted value for species definition but 97% or 99% are used in several papers.

The main purpose of this research was to define a standard procedure for 16SrRNA genes analysis.

A possible correlation between bacterial taxonomy and sequencing data was explored to understand how useful this analysis could be for species identification and searching a statistical improvement for this purpose. Starting from such data, similarity values within each taxonomic level have been defined and compared to understand if a pattern may subsist and if it could be applied to next-generation sequencing data interpretation.

A platform for data analysis was created starting from the definition of pairs of universal primers suitable for ultra high-throughput sequencing but also for less expansive procedures like RFLP.

These primers have been tested via bioinformatic analysis on sequenced ribosomal genes and on two samples offering several interesting features for microbiology-based industrial applications.

## Chapter 2. The Anammox bacteria

Water resources are currently identified as one of the major limiting factors for human development. A huge amount of industrial and domestic discarded water is directly poured into the environment causing increasing levels of pollution in a continuous process.

Ammonium deriving from chemical production (Van der Weerden et al. 1997), agriculture (Sutton et al., 2000), refineries, petrochemical, metallurgical processes (Vucebic, 1997) and wastewater treatment (Rostron, 2001) has become significant as environmental pollutant. Such molecules might lead to several negative effects on environment, included

eutrophication of lakes and rivers (Furer et al., 1996), on aquatic life (Balci et al., 2002) and may be responsible of several diseases.

Microbes contribute to the various transformations that nitrogen compounds undergo when are discarded into the environment. Ammonium can be converted to nitrite and the negative charge and high solubility of this molecule determine a transition from soil to the ground-water.

All these considerations, determined the implementation in 1991 of the Council Directive 91/676/EEC “*concerning the protection of waters against pollution caused by nitrates from agricultural sources*” (<http://eur-lex.europa.eu/>) and defining a procedure to limit the amount of nitrogen compounds issued by agricultural practices and prevent further pollution.

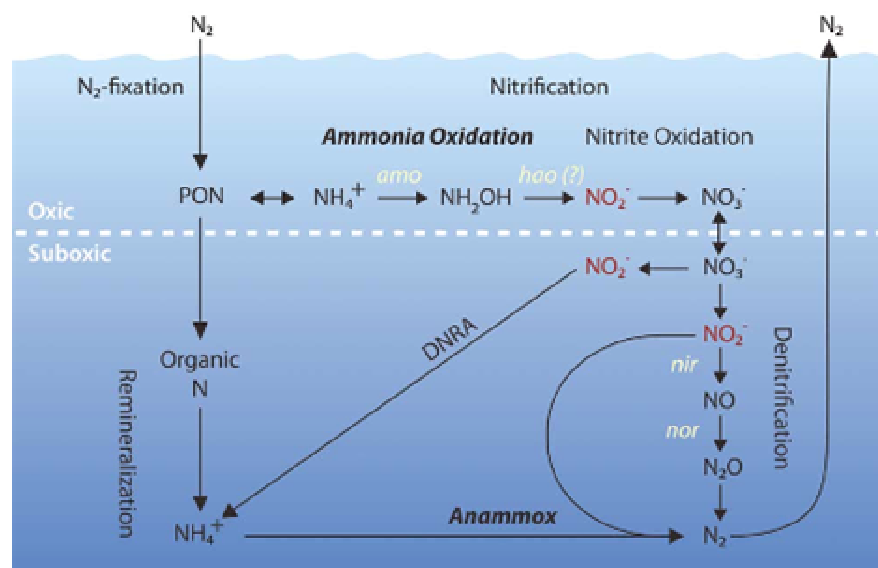
Efficient technologies that allow ammonium reduction are continuously developed and several applications require one or more specific bacterial activity to occur.

## 2.1 Nitrogen in nature

Nitrogen is an essential requirement for all organisms as a constituent of proteins and nucleic acids.

It is estimated that plant dry mass is made by ~2 to 5% of nitrogen. However the elementary dinitrogen gas which is one of the most abundant constituent of atmosphere can not be directly used by pluricellular organisms. At the basis of the biological nitrogen cycle, activity of diazotrophic bacteria plays a fundamental role to convert dinitrogen gas to nitrate which can be used by other organisms (Fig. 1, Francis et al., 2007).

Other bacterial species can catalyze the reverse reaction in which different form of nitrogen compounds oxidized to produce dinitrogen gas.



**Fig. 1:** Microbial nitrogen transformations above, below and across an oxic/anoxic interface in the marine environment. The largest reservoir of N is  $N_2$  gas and must be fixed by microorganisms before it is useable by other organisms. N exists in its most reduced state within organisms, but it is rapidly nitrified to nitrate aerobically when released following cell death and lysis. Nitrate is then denitrified to  $N_2$  gas under suboxic to anoxic conditions, completing the cycle. The anammox reaction has a central role in anaerobic nitrogen cycle. From Francis (2007).

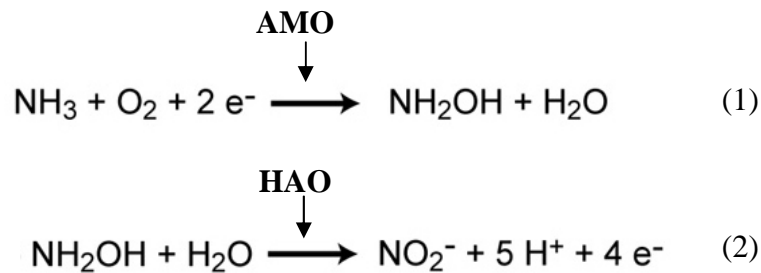
### 2.1.1 The nitrification process

Nitrification is a two step reaction that involves different bacterial species.

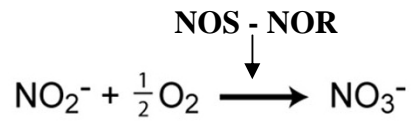
Obligate autotrophs *Nitrosomonas*, *Nitrospira*, *Nitrosococcus* and *Nitrosovibrio* are anaerobic chemolithoautotrophic ammonium-oxidizing bacteria that can use ammonium  $\text{NH}_4^+$  as energy source and molecular oxygen as electron acceptor.

Ammonia is firstly oxidized by the activity of the membrane-bound heterotrimeric copper enzymes ammonia monooxygenase (AMO) producing hydroxylamine  $\text{NH}_2\text{OH}$ .

A large fraction of hydroxylamine is oxidized by a periplasmic hydroxylamine oxidoreductase (HAO) while a small amount is oxidized by another periplasmic enzyme mono-heme cytochrome P460 protein (Arp et al., 2002; Norton et al., 2002; Bergman et al., 2003) (1 step of the reaction). Nitrite is produced by a combined action of such enzymes (Purkhold et al., 2000).



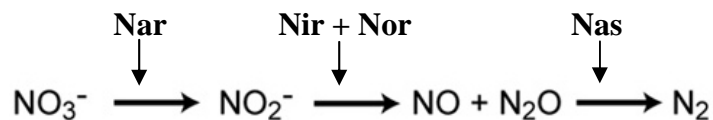
The other step of the reaction (2), determines the oxidation of nitrite to nitrate  $\text{NO}_3^-$ . This second-step reaction is carried out by bacteria belonging to *Nitrobacter*, *Nitrococcus* and *Nitrocystis* genera, phylogenetically poorly related and less studied if compared with ammonium oxidizing bacteria (Teske et al., 1994). Nitrite oxidoreductase NOR analyzed in *Nitrobacter* species (Starkenburger et al., 2006; Starkenburg et al., 2008) and nitrite-oxidizing system NOS identified in *Nitrospina*, *Nitrococcus* and *Nitrocystis* represent the key enzymes involved in nitrite oxidation (Spieck, Bergey's Manual® of Systematic Bacteriology).



Interestingly, nitrite represents the sole source of nitrogen and energy for species like *Nitrobacter*: because of this specific requirement, in natural environment it depends by *Nitrosomonas* or other nitrite-producers for nitrogen supplies (Moat et al., 2002).

### 2.1.2 The denitrification process

A different reaction, denitrification, consists in the reduction of oxidized nitrogen compounds as nitrite and nitrate to gaseous dinitrogen (Zumft et al., 1997).



This reaction is carried out by facultative heterotrophic bacteria that unlike other organisms, can use partially-oxidized forms of nitrogen as electron acceptors and poor organic matter as carbon and energy source. Several species are facultative denitrifiers and can use also O<sub>2</sub> as final electrons acceptor: in aerobic conditions, 686 kcal are produced by the complete oxidation of a single mole of glucose while 570 kcal are produced in anaerobic condition (Delwicke, 1970). Aerobic respiration represents a favored process and denitrification may occurs only if completely anoxic condition is maintained. The denitrification process, is activated when oxygen levels are low and nitrate becomes the primary oxygen source for microorganisms.

Four enzymes are involved in denitrification and are required for the complete

nitrite oxidation to dinitrogen gas (Moura et al., 2001). The complete enzymatic systems lacks in several bacterial species determining the production and accumulation of free intermediates (Kumar et al., 2010).

Nitrate reductase (*Nar*) is a membrane bound molybdenum-iron-sulphur protein which catalyzes the reduction of nitrate  $\text{NO}_3^-$  producing nitrite  $\text{NO}_2^-$  (Richardson et al., 2001). Oxygen may inhibit these enzymes.

The second enzyme is represented by a periplasmic copper- and heme- containing nitrite reductase (*Nir*) (Wherland et al., 2005) recognized only in denitrifying organisms which catalyzes the conversion of nitrite  $\text{NO}_2^-$  producing nitric oxide NO (Zumft et al., 1997; Nojiri et al., 2007). Starting from nitric oxide NO, nitrous oxide  $\text{N}_2\text{O}$  is produced due to the activity of a membrane bound nitric oxide reductase (*Nor*) (Hendriks et al., 2000; Hino et al., 2011). Finally, a periplasmic copper-containing protein, nitrous oxide reductase (*Nos*) determine the reduction of nitrous oxide producing  $\text{N}_2$  (Zumft et al., 1997; Zumft et al., 2005; Kumar et al., 2010).

Several species of denitrifying bacteria can be identified within Gram-negative Alpha- and Beta-Proteobacteria class. *Pseudomonas*, *Paracoccus*, *Alcaligenes* and *Thiobacillus* some examples of denitrifying bacteria belonging to these classes. Denitrifying activities have also been observed for some Gram-positive Firmicutes, in particular belonging to the *Bacillus* genus, and some halophilic Archaea (*Halobacterium* genus) (Ahn et al., 2006).

## 2.2 Biological approaches to nitrogen removal

Several strategies to lower the wastewater nitrogen content are currently applied involving chemical, physiochemical and biological methods. The best method to use relies on several considerations that include process cost-benefit, energy requirement, chemical and environmental sustainability. According to Mulder (Mulder et al., 2003), the ammonium concentration represents a fundamental parameter for for the appropriate strategy choice.

- $\text{NH}_4^+$  concentration lower than 100mg  $\text{NH}_4^+\text{-N/l}$ . In this range a biological N-removal should be preferred. Domestic wastewater is often comprised within this

limit.

- NH<sub>4</sub><sup>+</sup> concentration between 100-5000mg NH<sub>4</sub><sup>+</sup>-N/l. This range comprises industrial wastewater from pectin industry, landfill leachate, or tannery. For this amount of pollutants, a biological treatment can be preferred as well (Janus et al., 1997).

- NH<sub>4</sub><sup>+</sup> concentration greater than 5000mg NH<sub>4</sub><sup>+</sup>-N/l: for this industrial scale wastewater treatment, a physiochemical method must be preferred.

Several technologies can be applied in wastewater treatment plants and the ability of cooperating microorganisms to oxidize nitrogen compounds has been widely used for this purpose.

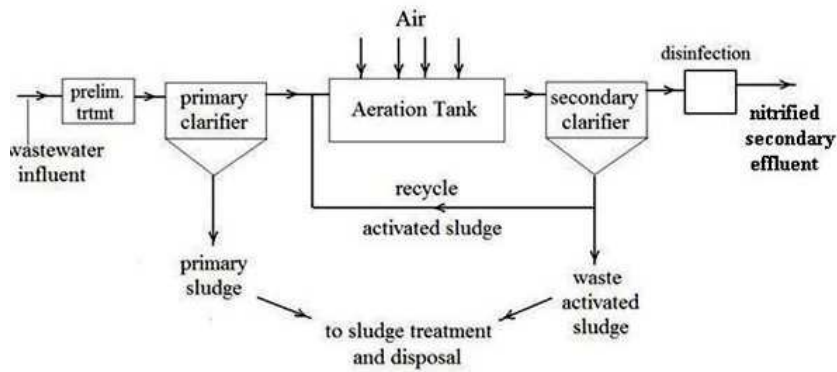
A typical treatment plant is represented by sequencing batch reactor (SBR). SBR are fill-and-draw based systems in which wastewater is added to the reactor, retained and treated by the sludge microfloral activity to remove pollutants and then discharged after sludge settling.

Nitrification and denitrification occur in the same reactor and require an alternation of oxygenated areas, where the nitrate is produced by chemoautolithotrophic aerobic bacteria, and anoxic zones in which nitrate is oxidized to gaseous nitrogen by facultative heterotrophic microorganisms..

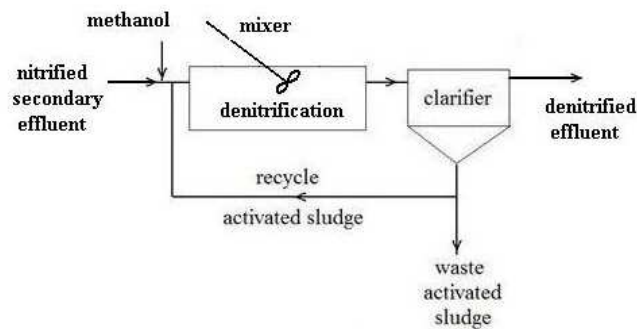
At end of the process treated water is removed leaving the reactive sludge inside the batch. This retention within the bioreactor facilitates the set up of the microbial consortium involved in the pollutant removal process.

Reaction requires time but the SBR can have a compact size reducing the space occupation and the operative sections for aeration, anoxic reaction and settling can be auto-regulated.

In the activated-sludge wastewater treatment process, the partial oxidation of nitrogen compounds occurs in different areas: an aerated zone which is fundamental for the activity of nitrifying bacteria (Fig. 2).



**Fig. 2:** first step of the activated-sludge wastewater treatment process. An aeration tank provides the optimal conditions for the aerobic ammonium oxidizer bacteria



**Fig. 3:** second step of the wastewater treatment. Denitrification can occur due to the anaerobic condition. Nitrogen gas represents the final product of reaction.

During the retention of wastewater in this compartment, ammonia nitrogen is oxidized to nitrate, due to nitrification and also COD are oxidized. Bacterial activity removes oxygen and creates a useful substrate that can be metabolized by denitrifying bacteria. A second, separated anaerobic environment characterizes these bioreactors: microbial activity requires the absence of oxygen and a carbon source (e.g. methanol) and determines the oxidation of nitrate to nitrite and oxidation of nitrite to dinitrogen gas which is given off to the atmosphere (Fig. 3). Coupled nitrification/denitrification systems, can remove more than the 95% of dissolved inorganic N.

### 2.3 The Anammox bacteria

ANAMMOX reaction (Anaerobic AMMonium OXidation) process was firstly discovered in 1995 (Mulder et al., 1995) but the reaction was previously predicted on the basis of thermodynamic evaluations in 1977 by Broda, to explain the existence of a single chemolithoautotrophic microorganisms able to convert ammonium to nitrogen gas and nitrite as chemical intermediate.



Indications of ANAMMOX activity in a denitrifying pilot plant emerged in 1999. Concurrents decreasing of ammonium and nitrate were observed while the nitrogen gas levels increased, suggesting a possible alternative to the known nitrification-denitrification process (Jetten et al., 1999)..

Moreover, the efficiency of anammox reaction appears to be unquestionably higher if compared with aerobic ammonia oxidation (Jetten et al., 2001a-b) even if it is slower (Fig. 4).

Parameter	Nitrification $\text{NH}_4^+ + \text{O}_2 \rightarrow \text{NO}_2^-$	Anammox $\text{NH}_4^+ + \text{NO}_2^- \rightarrow \text{N}_2$	Unit
Free energy	-275	-357	kJ/mol
Biomass yield	0.08	0.07	mol/mol C
Aerobic rate	200–600	0	nmol/min/ mg protein
Anaerobic rate	2	60	nmol/min/ mg protein
Growth rate	0.04	0.003	/h
Doubling time	0.73	10.6	days
$K_s \text{ NH}_4^+$	5–2600	5	$\mu\text{M}$
$K_s \text{ NO}_2^-$	N/A	<5	$\mu\text{M}$
$K_s \text{ O}_2$	10–50	N/A	$\mu\text{M}$

N/A, not applicable;  $K_s$ , affinity constant.

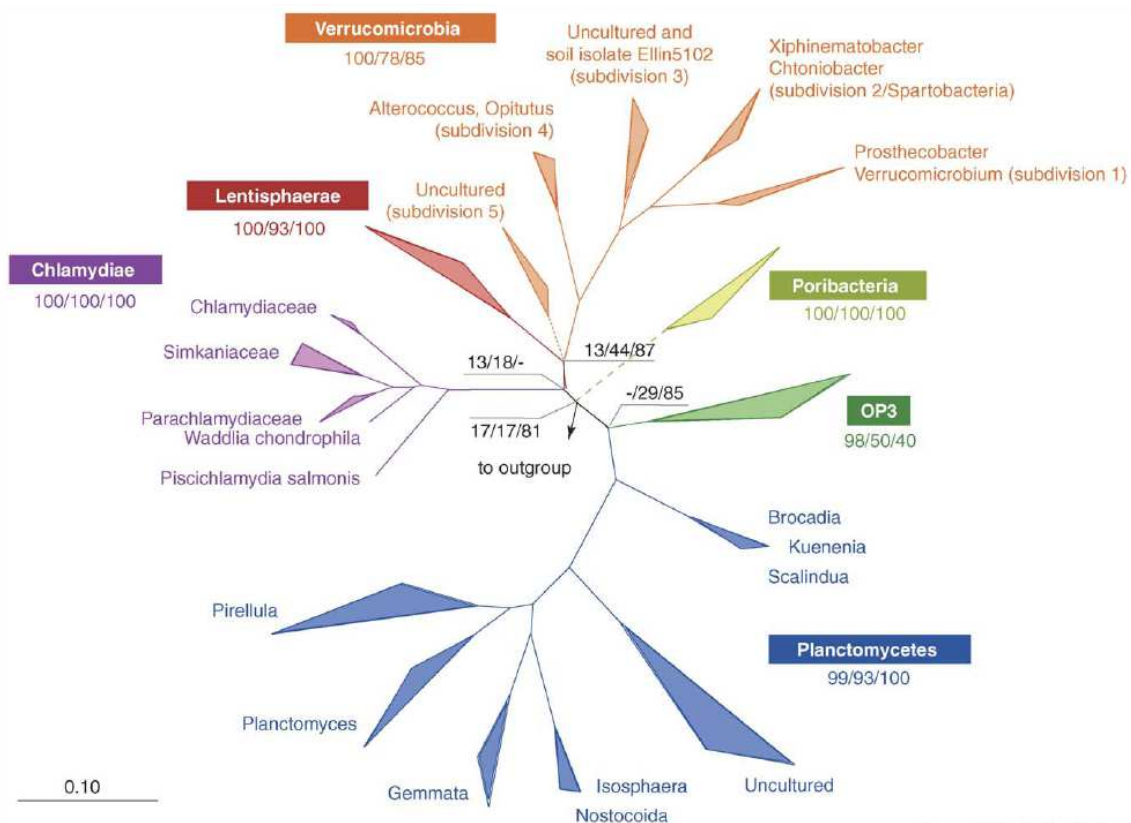
**Fig. 4:** Comparison and biochemical parameters of aerobic and anaerobic ammonia oxidation (Jetten, 2001)

## 2.4 A deep branch of prokaryotes: the Planctomycetes phylum

Currently identified Anammox bacteria are classified as members of Planctomycetes phylum which seems to occupy the deepest branch of the entire phylogenetic Bacteria tree (Brochier et al., 2002, Jun et al., 2010).

Planctomycetes are identified as members of a unique taxon comprising Planctomycetes, Verrucomicrobial, Chlamydiae and Lentisphaerae, Poribacteria and OP3 and defined the PVC superphylum (Fig. 5) (Wagner et al., 2006). Despite the differences that characterize organisms belonging to this group, an increasing number of genomic data, support the monophyly of the PVC superphylum.

Planctomycetes populate several environments as free-living species (Glöckner et al., 2003) or in association with other organisms (Hentschel et al., 2002, Shinzato et al., 2005).



**Fig. 5:** Phylogenetic relationships within the PVC superphylum based on a 16S rRNA gene comparative analysis. (Wagner 2006)

Several features characterize Planctomycetes, distinguishing them from other bacteria. One of the most most remarkable characteristic is represented by the presence of an internal membrane complex which defines a cytoplasmic compartmentalization. Two major regions can be surrounded by a single or a double bilayer membrane defining two compartments. The paryphoplasm is delimited by the cellular membrane and by the internal membrane complex: no ribosomes were detected in this region (Lindsay et al., 2001) but the presence of RNA underlines the existence of this structure as a true compartment of cytoplasm (Lindsay et al., 1997). Pirellulosome, defined also as riboplasm or ribosome-containing-cytoplasm, represent the standard cytoplasmic compartment where ribosomes and the nucleoid are situated (Jetten et al., 1999).

Interestingly, Planctomycetes differs from other Bacteria in the organization and composition of the layers that surround cells. Although, they lack in peptidoglycan as a cell wall component (van Niftrik et al., 2004) and genomic analysis revealed the presence of several ancestral genes involved in this structure biosynthesis (Strous et al., 2006). Moreover, unlike other Gram-negative bacteria, Planctomycetes possess a cell wall that is not surrounded by any membrane (van Niftrik et al., 2010).

#### **2.4.1 Anammox bacteria: morphology, physiology, genomics**

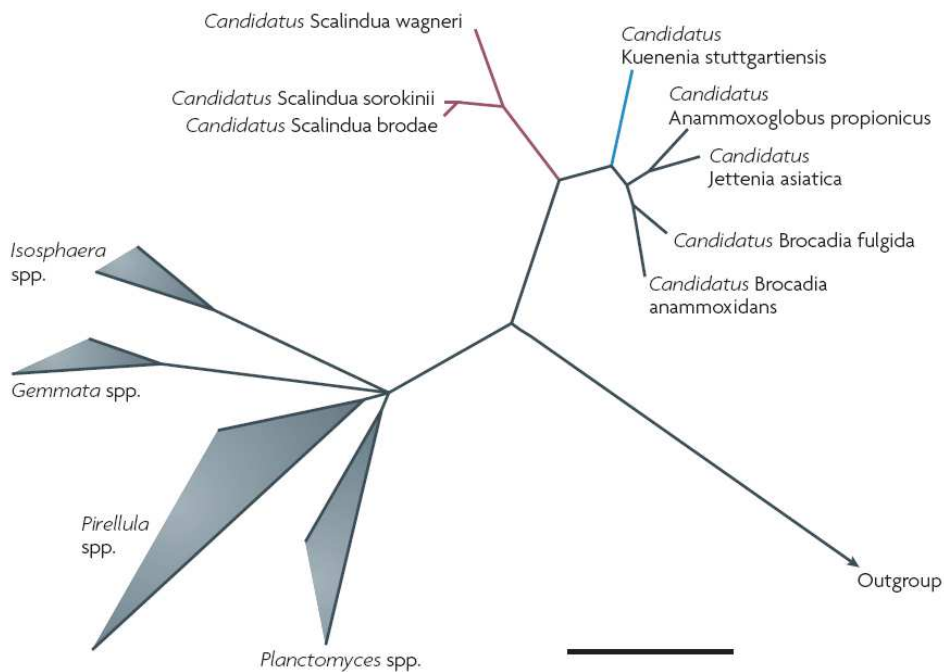
Currently, five genus of Anammox bacteria were identified: *Brocadia* (Mulder et al., 1995) and *Kuenenia* (Schmid et al., 2000) were isolated from bioreactors set up starting from biomasses isolated from wastewater treatment plants. *Anammoxoglobus* was firstly identified in a bioreactor: this anammox bacteria can oxidize propionate in the presence of ammonium, nitrite and nitrate (Kartal et al., 2007). *Jettenia* was found in a granular sludge anammox reactor (Quan et al., 2008).

*Scalindua* was discovered in a wastewater treatment plant but also identified in low oxigenic marine zones and anoxic sediments (Schmid et al., 2003).

Anammox bacteria play an important role in the natural nitrogen cycle.

It is estimated that Anammox microorganisms activity is responsible for at least 50% of the removal of fixed nitrogen and re-emission of dinitrogen gas in the atmosphere.

They were detected in several natural environments, marine sediments (Dalsgaard et al., 2002), in deep sea hydrothermal vents (Byrne et al., 2009), Black sea (Kuypers et al., 2003) and other anoxic zones that comprehend marine and fresh water environments and iced sea and permafrost soil as well (Fig. 6) (Kuenen et al., 1998; Francis et al., 2007).

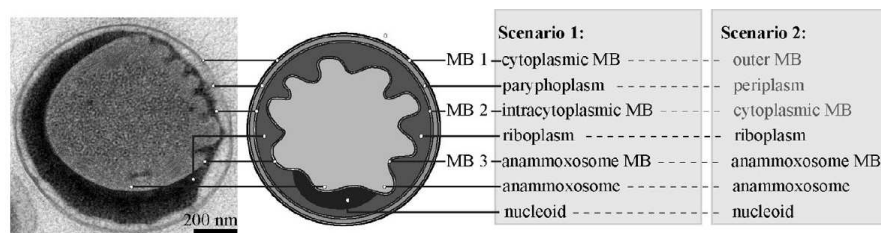


**Fig. 6:** 16S rRNA gene-based phylogenetic relationship of the different families of anammox bacteria. The monophyly of Anammox bacteria within the Planctomycetes phylum is shown. The scale bar represents 10% of sequence divergences (Kuenen, 1998)

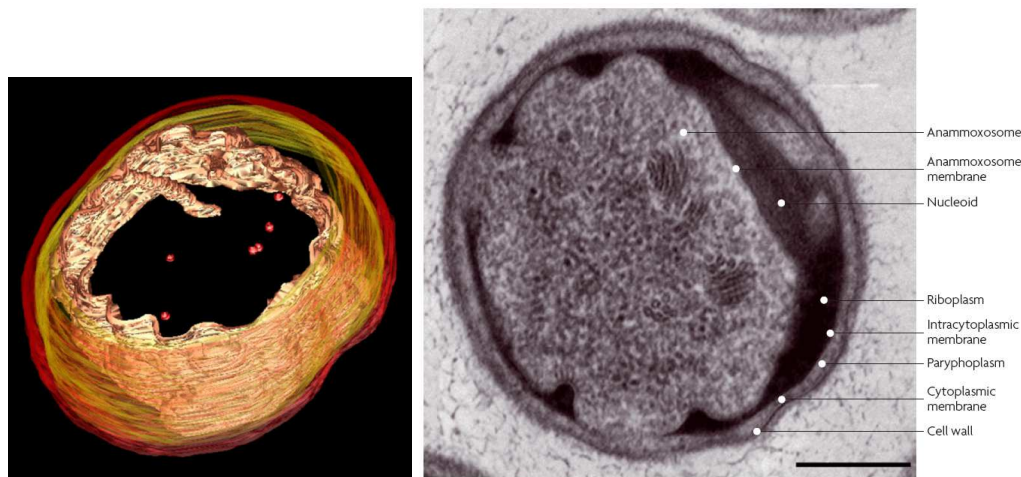
Anammox bacteria are coccoid chemolithoautotrophic cells with a diameter of less than 1  $\mu$ m. A typical red color, budding production and crateriform structures on cell-surface are their main characteristic (Zhang et al., 2008). Their doubling time

is comprised between 10 and 30 days and they can survive in natural environments in a range of temperature that starts from  $-2.5^{\circ}\text{C}$  in sea ice (Dalsgaard et al., 2002; Rysgaard et al., 2004) and can reach  $70^{\circ}\text{C}$  in hot spring hydrothermal vent zones (Byrne et al., 2009; Jaeschke et al., 2009). On the other hand, Anammox microorganisms isolated from wastewater treatment plants can grow in a range of temperatures between  $20^{\circ}\text{C}$  and  $43^{\circ}\text{C}$  with an optimum around  $39^{\circ}\text{C}$  and pH comprised between 6 and 8.9. Oxygen represents an inhibitor of Anammox activity, confirming the anaerobic nature of these organisms: this action is completely reversible and when all oxygen traces are removed the anammox reaction takes up again.

A particular intracellular compartment characterizes Anammox bacteria: it is called anammoxosome and represents the structure in which anammox reaction occurs catalyzed by several enzymes that are localized here. This compartment, takes up the 30% of the total cell volume (Van Niftrik et al., 2004) and it is involved in generating a proton motive force across the membrane which is used to synthesize ATP (Fig. 7). This intracellular compartmentalization prevents the diffusion of protons and other toxic reaction intermediates across the cell (Lindsay et al., 2001).



**Fig. 7:** Ultrastructure of an Anammox bacterium and the different scenarios concerning its cell plan MB, membrane (From van Niftrik 2010).

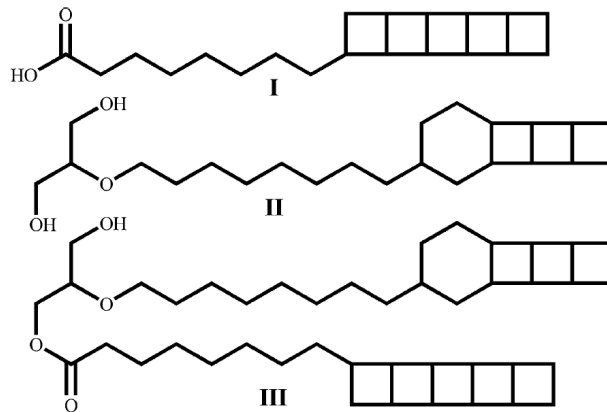


**Fig. 8:** electron tomography on *Candidatus Kuenenia stuttgartiensis* (left). From the outside to the inside, cell wall (red), intracytoplasmic membrane (yellow) and anammoxosome membrane (pink) are visible (from van Niftrik, 2008).

Characteristic structures of *K. stuttgartiensis* are displayed using transmission electron microscopy (from Kuenen, 2008).

This organelle seems to be completely isolated from the other cellular membranes, unlike other Planctomycetes in which a continuing structure is constituted by internal membrane complex and paryphoplasm membranes (Fig. 8).

Anammoxosome membranes, are constituted by a peculiar class of lipids (Sinninghe Damsté et al., 2002): C18 and C20 fatty acids chains containing 3 or 5 concatenated cyclobutane rings bounded to a glycerol backbone or to an alkyl chain and defined 'ladderanes' (Fig. 9) (Sinninghe Damsté et al., 2005). These lipids form an extremely dense structure that forms a physical obstacle to diffusion of molecules and intermediates. Thanks to this feature, Anammox cells are able to maintain a stable concentration gradient despite their slow metabolism and are able to protect all their cellular components from toxic reaction intermediates such as hydrazine (Hopmans et al., 2006). Ladderane lipids were detected only in this bacterial taxon and represent an useful molecular marker for the environmental Anammox detection and identification (Sinninghe Damsté et al., 2002).



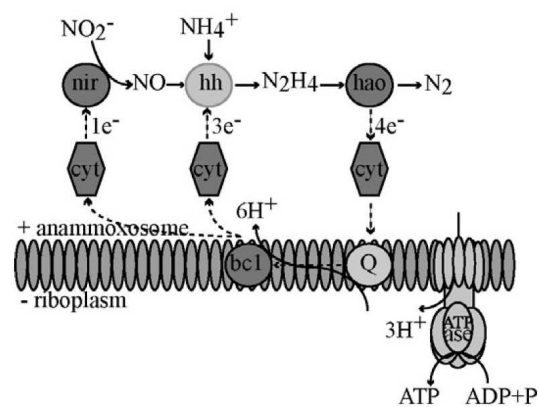
**Fig. 9:** structure of ladderane lipids (From Jetten et al., 2003)

## 2.4.2 Biochemistry of the Anammox reaction

A complete model for the anammox reaction in which biochemical data and gene content were considered simultaneously, was hypothesized.

After the analysis of the anammox bacteria *Kuenenia stuttgartiensis* genome annotation (Strous et al., 2006; ), several modifications of the previously hypothesized metabolic pathway have been introduced, defining the present model that explains in the best way the anammox reaction (Fig. 10) (van Niftrik et al., 2010).

The biochemical model sees the Anammox reaction catalyzed by several cytochrome c related proteins.



**Fig. 10:** proposed model for the Anammox reaction and coupling with ATP synthesis. hh, hydrazine-hydroxylamine oxidoreductase; Hh, hydrazine idrolase; nir, nitrite reductase; bc1, citrocrome bc1 complex; cyt, cyocrome; Q, co-enzyme-Q. (From val Niftrik et al., 2010)

Nitrite ( $\text{NO}_2^-$ ) is firstly reduced to nitric oxide (NO) due to the activity of a nitrite reductase NirS encoded by the *nir* genes. Nitrite reduction was not predicted by previous models but this gene annotation suggests a possible NO production as reaction intermediate. Nitrite is then combined with ammonium ( $\text{NH}_4^+$ ) by a hydrazine hydrolase enzyme (hh) to form hydrazine ( $\text{N}_2\text{H}_4$ ). Finally, hydrazine is oxidized to dinitrogen gas ( $\text{N}_2$ ) by a hydrazine-hydroxylamine oxidoreductase (HAO) which represents the key enzyme for the anammox reaction in anammoxosome and constitutes more than 10-15% of total cell proteins (Shalck et al., 2000).

The four electrons derived from hydrazine oxidative processes, enter a respiratory chain at ubiquinone level.

In agreement with the current model, anammox reaction determines a proton gradient due to the translocation of proton from the cytoplasm to the anammoxosome. The differential between the internal and external sides of the anammoxosome membrane, produces a proton-motive-force used by membrane-bound ATP-ases to produce ATP (van Niftrik et al., 2010).

### **2.4.3 Potential application of the Anammox process**

Nowadays the current experimental technologies for a large scale application of the anammox reaction seem to indicate that a higher efficiency can be achieved only if an Anammox and denitrifiers is established.

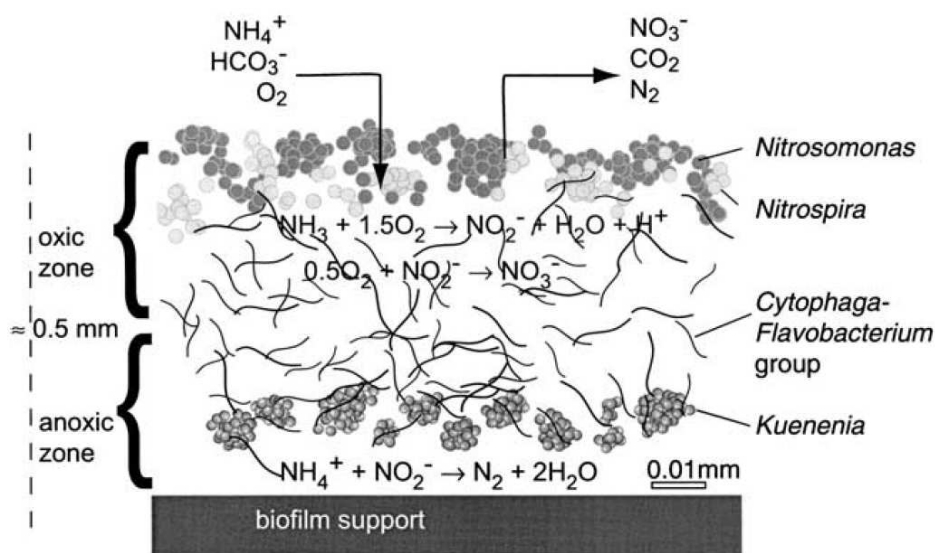
The process can be implemented applying several technology combination. Two different steps can be separately carried out to selectively obtain a functional Anammox-type consortium: in the OLAND process (Oxygen-Limited Autotrophic Nitrification-Denitrification), during the first step, a biocatalyst consisting in a community of aerobic ammonium-oxidizers bacteria is used. Ammonium- and nitrite- rich environments obtained after this first step represent the optimal substrate for Anammox bacteria to operate (Kuai et al., 1998).

In the CANON process (Completely Autotrophic Nitrogen removal Over Nitrite), the two reactions begin at the same time: aerobic ammonium oxidizer bacteria,

under oxygen limitation, oxidize ammonium to nitrite which is consumed by anammox bacteria to give dinitrogen gas. Moreover, they remove all oxygen traces ensuring the complete microenvironment anaerobiosis essential for any Anammox activity (Slijkers et al., 2002).

SNAD bioreactors (Simultaneous partial Nitrification-Anammox-Denitrification) couple CANON process with the activity of denitrifying bacteria converting ammonium to nitrogen gas and dissolved organic carbon to carbon dioxide (Chen et al., 2009).

Using different molecular biology techniques, the interactions and spatial distribution of the Anammox-community bacteria was analyzed (Fig. 11).



**Fig. 11:** spatial distribution of bacterial species in a CANON biofilm (from Egli, 2003)

Nitrifying bacteria such as *Nitrosomonas* and *Nitrospira*, colonize the superior oxygenated layer, consuming oxygen and producing  $\text{CO}_2$  and nitrite. Nitrifying bacteria provide environment and substrates for the anammox bacteria activity (Hao et al., 2002; Egli et al., 2003).

#### **2.4.4 Pros and cons of anammox reactions in industrial applications**

Anammox bacteria represent an interesting alternative solution for the development of bioreactors for the treatment of nitrogen-rich wastewater.

The entire oxidation of ammonium is carried out by a single species while in the denitrification-nitrification solution at least two species are required. This, implies that once isolated and enriched the Anammox population, the process can proceed autonomously and the optimal conditions must be maintained for a single bacterial species. Usage of nitrifying bacteria and denitrifying bacteria in a single community (e.g. SBR) requires that optimal conditions for each species must be maintained and monitored taking in consideration that microbial growth must follow the establishment of a delicate equilibrium between species.

Usage of different tanks for the two-step treatment, in which aerobic ammonium-oxidizer and denitrifying bacteria are located in two distinct compartments, requires open spaces and a complex machinery for the waste displacement from one chamber to another and for substances recovery and transfer from one chamber to another. Anaerobic conditions required by Anammox bacteria do not require a dedicated plant for oxygenation which is needed for nitrifying bacteria.

Finally, since Anammox bacteria are autotrophs, they do not require external carbon source if the optimal growth conditions are maintained. Denitrifying bacteria need rapidly metabolizable organic sludges or an external carbon source like methanol for optimal growth (Egli et al., 2003).

It is estimated that application of Anammox bacteria to industrial scale bioreactors, could reduce power consumption by up to 60%. A reduction of CO<sub>2</sub> emissions by up to the 90% and reducing excess sludges, an Anammox bioreactor requires up to 50% less spaces if compared to conventional denitrification/nitrification systems (from [www.paques.nl](http://www.paques.nl)).

Advantages are accompanied by disadvantages. One over all is represented by the long doubling time of Anammox bacteria (~11 d) and the low rate of biomass production that determine long time for the start-up of the bioreactors.

Several substrates can have negative effects on Anammox bacteria: acetylene and phosphates and oxygen which may enter in the system may inhibit the reaction

(van de Graaf et al., 1995): concentrations greater than 1mM, and 2□M and between 5 and 10mM of phosphate, oxygen and nitrite, determine a reversible inactivity of Anammox bacteria (Jetten et al., 2001b) and probably high ammonium concentrations may have deleterious effects (Waki et al., 2007). Other molecules, glucose, formate and acetate do not seem to affect the Anammox activity (Kartal et al., 2007) while an interesting role was described for propionate. This molecule, represents an energy source for the nitrite to nitrate reduction (Guyen et al., 2004; Guven et al., 2005) but the mechanisms of reaction remain unknown. Propionate oxidation occurs simultaneously with ammonium oxidation (Op den Camp et al., 2006) as observed for *Anammoxoglobus propionicus*.

## **2.5 Aim of the research**

The present research investigates biomasses collected from a pilot bioreactor developed by Eurotec Water Treatment Technology Company of Padua, in which Anammox activity was observed, using both classical molecular biology techniques and next-generation sequencing.

This bacterial community was selected from a nitrification-denitrification community deriving from a wastewater treatment plant that collects and process organic residues from poultry farms. The creation of an Anammox-type consortium was promote providing an appropriate nourishing solution made of nitrite, ammonia and microelements in a temperature controlled system.

The biological question was focused on verifying if the observed Anammox activity observed would be related to the presence of already known Anammox bacteria.

The next generation sequencer, with its ultra high-throughput data production represents the most powerful technology to detect and identify microorganisms.

Moreover, the general composition of this bacterial community was monitored in time, starting from initial inoculum until the reaching of a functional state, with the aim of identify which bacterial genera was the most representative and to determine the effects of different growth conditions in promoting the selection of

any particular microorganism.

The development of a tool for bioreactors management was then carried out. Using a RFLP based approach, the definition of a restriction pattern for the Anammox-community was determined for each stage of the consortium development and a 'standard pattern' useful for bioreactor monitoring and maintenance was obtained. Taking into account both sequencing data and resulting analysis, a functional microbial community was individuated and a tool for monitoring its global status was devised. Focusing on sequencing data, the definition of a 'standard' composition for the bioreactor community can be helpful to screen different and more efficient inocula composition to promote the development of a functional and stable community.

Biomasses collected from the EurotecWTT pilot bioreactor were used by the group of Professor F. Adani (University of Milan) to set up another pilot batch to increase rate and efficiency of the anammox reaction. Using the biomasses collected from this pilot batch and from EurotecWTT bioreactor, several distinct reactor batches were set up. Different growth conditions and nourishing solutions were tested and biochemical data, ammonium and nitrite were monitored to optimize growth rate and anammox reaction efficiency.

Once obtained a global view concerning bacterial pathways selection and biochemical data, ten samples were selected for a 16S rRNA genes tag-sequencing by Roche-454. Bioinformatic tools for sequences annotation and analysis were developed to find a correlation between different data types.

### **Chapter 3. The biological system of the dromedary rumen**

In ruminants, rumen represents the anatomical structure in which the hydrolysis of vegetable fibers and part of nutrients absorption occurs.

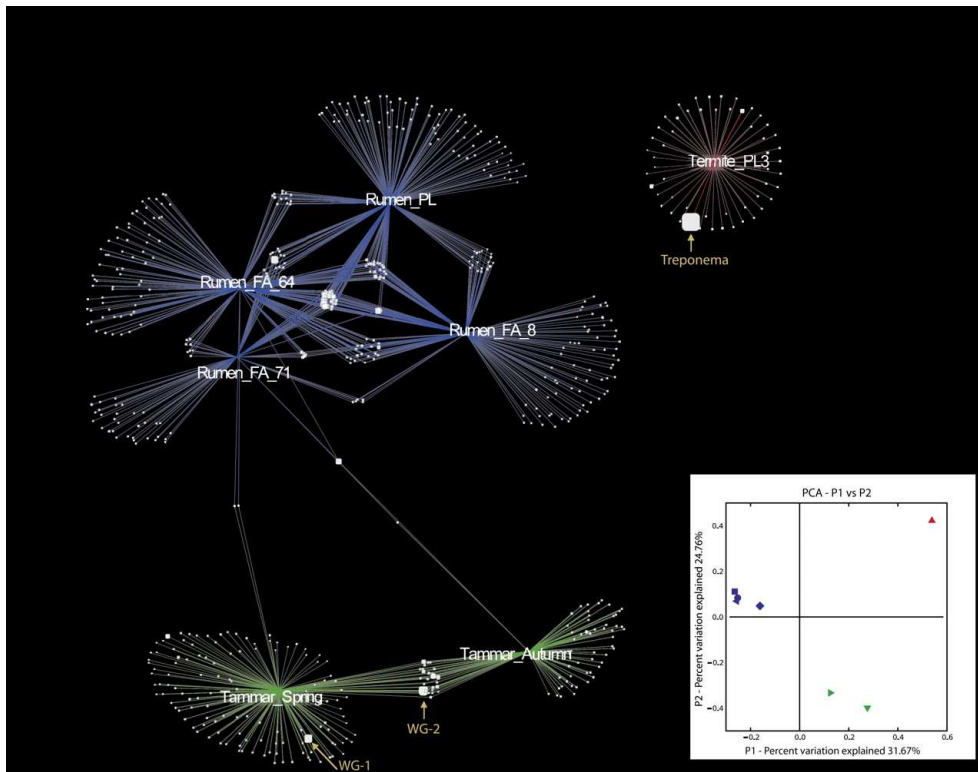
The ability to derive energy from complex substrates represented by the plant cell wall, is the consequence of the specific activity of the ruminal microflora. Such consortium is constituted by a community in which a total of  $10^{10}$ /ml of at least 30 predominant bacterial species,  $10^6$ /ml cells of 40 species of protozoa and  $10^5$ /ml cells of several species of fungi collaborate to the digestive processes.

Analyzing the high specificity of the genes and the number of individual cells, prokaryotes can be considered the most important players and would justify the success of fibers degradation and utilization during the entire processes (Miron et al., 2001).

The symbiotic association with a rich microflora characterizes the digestive tract of all the living organisms, invertebrate and vertebrate (from termite to humans, Barcenilla et al., 2000; Shinizato et al., 2005, Warnecke et al., 2007) and it is fundamental for life. While in the carnivores and omnivores, bacteria are mainly involved in vitamins production in all the herbivores play a central role in food degradation and synthesis of useful molecules as volatile fatty acids involved in energy production (Hungate, 1984)..

Using culture-dependent analysis and new technologies related to high throughput sequencing, the microscopic world of several ruminants has been analyzed. In particular 16S rRNA genes analysis have allowed a more complete and deep description of rumen microbiota, revealing the extreme complexity and the huge number of microorganisms that remain uncharacterized and uncultured (Tajima et al., 2001). Whole metagenomics and metatranscriptomics data have defined the several enzymatic pathways and activity which represent the basis of this interaction.

A clear phylogenetic relationship, can be observed comparing microbial populations and their hosts: poorly related organisms exhibit a poorly related microflora while closely related animals share an high number of bacterial species (Fig. 1- Pope, 2011).



**Fig. 1.** OTU comparison between 16S rRNA genes sequenced from the Tammar wallaby (green), bovine (blue) and termite (red). The cellulolysis is a common process that characterizes all these digestive tracts but a close relationship between microbiomes and host is clearly showed (Pope, 2010).

Different bacterial genomes and genes, highlight a different evolution and metabolic pathways to perform a common digestive process (Pope et al., 2010).

### 3.1 The ruminal microbiota

Ruminal microflora represents a lignocellulosic and fibers degradation bioreactor for the conversion of the plant wall to biomass for the organism. This chemical process requires the cooperation and a close interaction between different species (Chen et.al, 2001).

In light of the newest knowledge about such environment, several relationships have been defined between bacterial community and ingested food, describing the rapid species variation as response to the different substrates that can be digested.

Despite the considerations of such flexible community, it is possible to define a “

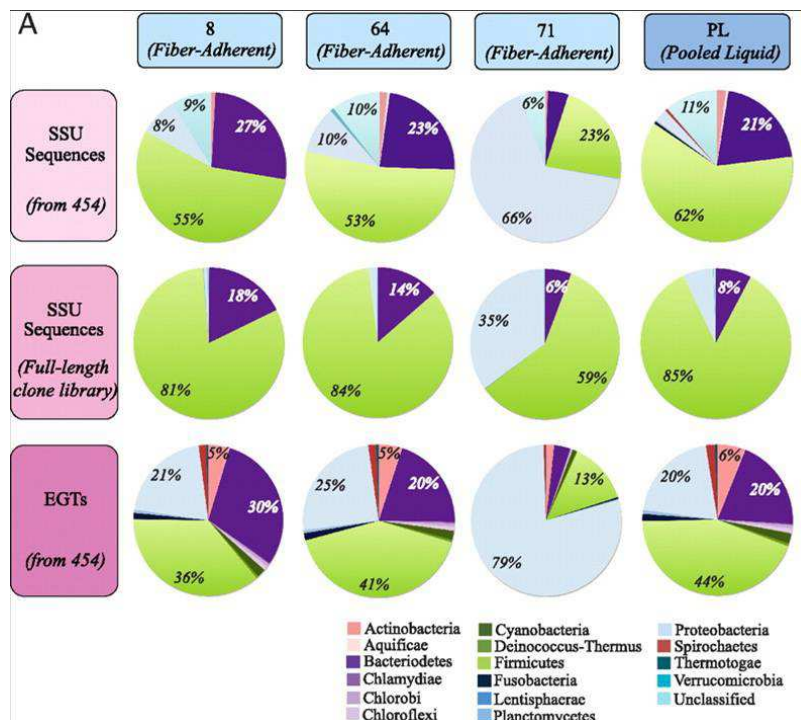
core” of bacterial species which may be considered the “basic community” that characterize the digestive apparatus of different ruminants. Specific role and functions were deciphered for some of these. In agreement with the description reported by several authors (Cheng et al., 1977, Czerkawski., 1988) ruminal bacteria can be identified considering the specific environmental existence:

- free living bacteria associated with the rumen liquid phase
- bacteria poorly associated with feed particles
- bacteria closely related to the feed particles
- bacteria associated with the animal epithelium
- bacteria in relation with protozoa and fungal sporangia

Bacterial species and activity discovered in every ruminal environment, reflect the high specialization level and the huge variability of this community. The liquid phase is characterized by non-cellulosolytic species belonging to Proteobacteria phylum and Tenericutes phylum where the solid particles associated microflora is represented by several cellulosolytic Clostridium class-related organisms as *Ruminococcus flavefaciens* and *Eubacterium siraeum*, and *Prevotella* species as ubiquitous clade (Tajima et al., 1999).

Within the food particle-associated bacteria, several differences can be observed considering poorly-associated and closely-related species, and the biodiversity is higher than in planktonic phase (Larue et al., 2005).

A consideration could be formulated observing the specific taxa distribution: during the digestive processes, bacterial cells associated to the solid particles are digested, constituting a huge supplementary source of nutrients for the animal. Ruminal microflora decreases during every nutritional cycle. In light of this, all the involved species should be always present as a “baseline” in the lumen cavity of rumen. In this direction, a valuable result was reported by Brulc et al (PNAS, 2009).



**Fig. 2.** Comparison of taxonomic variation in the rumen of three animals (8, 64, 71) fed with the same diet. Data were obtained by 16S rRNA genes 454 pyrosequencing, 16S rRNA genes come from full length clone libraries and environmental gene tags (EGTs). (Brulc, 2009).

In Fig. 2, fiber adherent and liquid phase taxa were identified: with the exception of the sample number 71, a comparable proportion between the principal phyla was observed.

Taxa distribution among rumen components reflects the different enzymatic activity of every single species. In particular, activity of polysaccharidase and glycosidases involved in polysaccharides degradation, is higher for ruminal solid phase, representing from the 88% to the 91% of the total cellulolytic activity (Miron et al., 2001). In the liquid phase, these enzymes are characterized by a lower activity which is comparable with other unicellular organisms (like protozoal) (Michalet-Doreau et al., 2001).

The largest contribution to the digestive process seems to be determined by feed particle related bacteria due to an higher cellulolytic activity. Generating a biological and biotechnological interest, they are largely studied. Studies focused on free living bacteria, while epithelium-associated and eucaryotes-associated remain rarefied.

### 3.2 Ruminant bacteria: *Fibrobacter*, *Ruminococcus* and *Butyrivibrio*

Classical methods (Gram staining) and molecular approaches (16S rRNA genes comparison) have been used to identify ruminal bacteria. *Fibrobacter succinogenes*, *Ruminococcus albus* and *Ruminococcus flavefaciens* represent the most abundant cellulolytic species recognized in distantly related ruminants (as reported below), while *Butyrivibrio proteoclasticus* and *Prevotella* represent the most abundant xylanolytic bacteria (Flint et al., 2008).

Actually, a wide number of the enzymes involved in the cellulolytic process has been detected and clarified and bacteria that characterize rumen can be grouped if capability and mechanisms of action during cellulose, starch and other type of fibers hydrolysis are compared.

#### 3.2.1 *Fibrobacter* sp.

*Fibrobacter* represents a genus of rod-shaped non-motile and strictly anaerobic Gram negative bacteria initially classified within Bacteroidetes phylum, genus Bacteroides, but now resolved in the distinct phylum of Fibrobacteres (Hungate, 1950). Only two species located in the mammalian gastrointestinal tract have been identified: *Fibrobacter intestinalis* and *Fibrobacter succinogenes*.

*Fibrobacter succinogenes* represents the first and major cellulolytic bacterium isolated from cattle and sheep rumen using cultural methods and confirmed as predominant by molecular approaches (Tajima et al., 2001, Jun et al., 2007a).

*Fibrobacter intestinalis* was firstly isolated from rat and found in bovines and monogastric animals. Less than 20% of total DNA homology, indicates the distant phylogenetic relationship between these organisms (Amman et al., 1992, Qi et al., 2008).

Four distinct groups can be identified for *Fibrobacter succinogenes* and the 16S rRNA genes similarity is between 95.3% and 98.1%, reflecting an intraspecific heterogeneity. High levels of metabolic activities without a specific preference for any plant tissues, indicate *group 1* as a major contributor involved in fiber digestion ( Kobayashi et al., 2008).

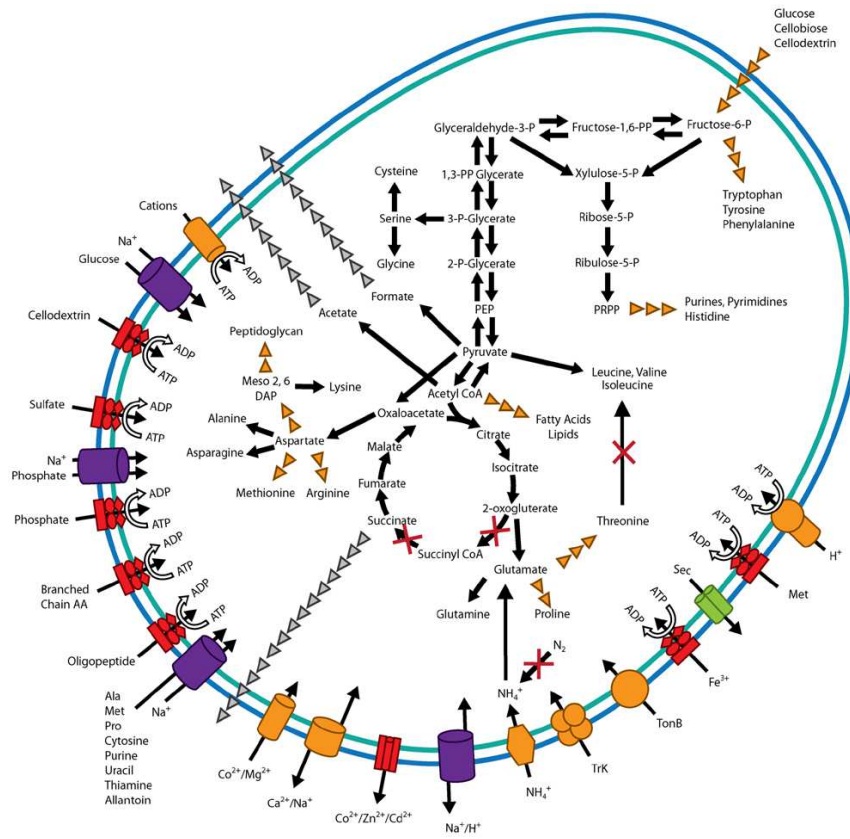
An higher cellulolytic activity can be observed for *F. succinogenes* if compared with the other fibrolytic bacteria. A closer association with the undamaged plant tissues characterizes this species whereas other ruminal bacteria adhere at level of inner layers and cell wall openings (Shinkai et al., 2007). *Fibrobacter* can be active on less digestive fibers and tissues and a large amount of such bacterium was recognized in animals fed by low quality forage rations.

Several enzymes involved in hydrolysis of polysaccharides were annotated in the *Fibrobacter succinogenes* genomes: the ability of hydrolyze a variety of complex sugars is coupled with the utilization of cellulose and its hydrolytic products as reported by growth assays (Suen et al., 2011). As observed for other cellulolytic bacteria, *Fibrobacter* does not excrete any cellulolytic enzyme and its activity requires the close interaction between bacterial cells and substrate.

Absence of a surface-bound structure for the substrate hydrolysis, cellulosome and related signature (as scaffoldins or dockerin-binding domains), distinguish *Fibrobacter* from the other cellulolytic bacteria. Several identified proteins may have a role during the initial interaction between bacterial outer membranes and substrate (Jun et al., 2007b) but the entire process, from adhesion to cellulose hydrolysis remain unclear (Brum et al., 2011, Suen et al., 2011).

This bacterium can derive its energy only from cellulose, degrading all allomorph molecules, including cellulose II (Weimer et al., 1991). A huge number of enzymes which can potentially degrade several types of polysaccharide have been identified

Considering the genome annotation, main pathways that characterized this organism were hypotesized (Fig. 3, Suen, 2011).



**Fig. 3.** Metabolic reconstruction based on genome annotation data. Missing enzymes from metabolic pathways are indicated with a red cross. *F. succinogenes* does not have an Entner-Doudoroff pathway or a glyoxylate shunt and has an incomplete pathway for the utilization of galactose, mannose, fructose and pentose sugars, confirming the inability to grow on any other substrate than cellulose (Suen, 2011).

Cellodextrin produced by the cellulose hydrolysis, can be imported into the cytoplasm due to a specific transporter. Inside the cell, this molecule is converted to glucose-1-phosphate by a cellodextrin phosphorylase.

A complete Embden-Meyerhof-Parnas pathway and an incomplete citric acid cycle without alpha-ketoglutarate dehydrogenase and a succinil-CoA-synthase (Miller, 1978) were identified in *Fibrobacter*, determining a succinate production as the major fermentative end products. Acetyl-CoA produced by a formate C-acetyltransferase, would represent the other volatile fatty acids synthesized.

Interestingly, a xylanase activity higher with respect to other species as *Ruminococcus* (Saluzzi et al., 1993), determines the conversion of xylan to xylose but *Fibrobacter* can not utilize this metabolic pathway as carbon source because of the lack of other related genes and proteins.

### 3.2.2 *Ruminococcus* sp.

*Ruminococcus* genus is represented by Gram-positive, obligate anaerobic organisms belonging to the Firmicutes phylum, class Clostridia, isolated for the first time in 1951 by Sijpestein (Sijpestein, 1951) and actually comprises 9 species (ncbi.nlm.nih.gov/Taxonomy/, online source).

With *Fibrobacter succinogenes*, *Ruminococcus* species as *Ruminococcus flavefaciens* and *Ruminococcus albus* represent some of the most abundant and active cellulolytic bacteria that characterize the ruminant microflora.

As observed for *Fibrobacter*, a close interaction between bacterial cells and substrate represents a fundamental step for the hydrolytic activity but several differences concerning such mechanism were found .

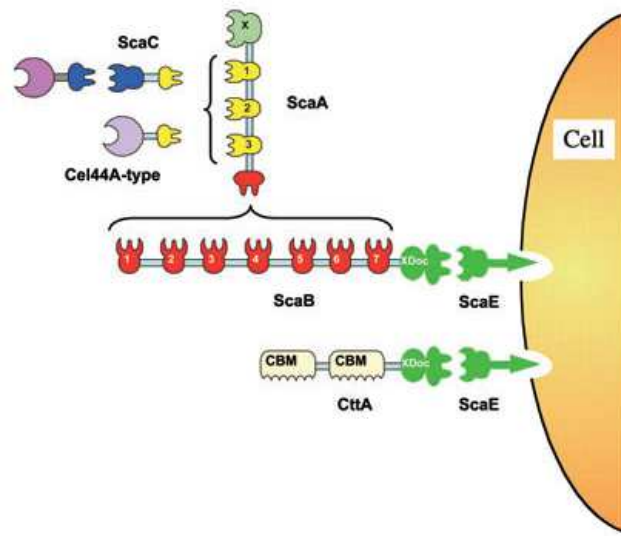
The activity of an high molecular weigh surface complex, defined cellulosome, allows *Ruminococcus* species to digest several cellulose types (Doerner et al. 1990, Miller et al., 2009)..

The cellulosome, was identified for the first time in 1983 in *Clostridium thermocellum* (Lamed et al., 1983). This complex, has been identified in anaerobic bacteria and fungi (Fig. 4, Doi et al., 2004) and represent a remarkable structure on the cells surface.

Microorganism	M/T*	Source
<b>Anaerobic bacteria</b>		
<i>Acetivibrio cellulolyticus</i>	M	Sewage
<i>Bacteroides cellulosolvens</i>	M	Sewage
<i>Butyrivibrio fibrisolvens</i>	M	Rumen
<i>Clostridium acetobutylicum</i>	M	Soil
<i>Clostridium cellobioperum</i>	M	Rumen
<i>Clostridium cellulolyticum</i>	M	Compost
<i>Clostridium cellulovorans</i>	M	Wood fermenter
<i>Clostridium josui</i>	M	Compost
<i>Clostridium papyrosolvens</i>	M	Paper mill
<i>Clostridium thermocellum</i>	T	Sewage soil
<i>Ruminococcus albus</i>	M	Rumen
<i>Ruminococcus flavefaciens</i>	M	Rumen
<b>Anaerobic fungi<sup>‡</sup></b>		
<i>Neocallimastix patriciarum</i>	M	Rumen
<i>Orpinomyces jayonii</i>	M	Rumen
<i>Orpinomyces PC-2</i>	M	Rumen
<i>Piromyces equi</i>	M	Rumen
<i>Piromyces E2</i>	M	Faeces

**Fig. 4.** The cellulosomes identified in Bacteria and fungi. M=Mesophilic, T= Thermophilic. (Doi, 2004)

Despite the differences related to the structure, proteins and aminoacidic sequences, cellulosomes of different organisms display several common features. A close interaction between fibrillar scaffolding proteins, scaffoldins, and enzyme subunits determines the tridimensional structure of the cellulosome: a scaffolding binding site, cohesin, and an enzyme cohesin-binding site, dockerin, represent the junction points between proteins. *R. flavefaciens* and *R. albus*, belong to the same genus but some differences characterize their cellulosome.



**Fig. 5.** Schematic Representation of the cellulosome complex of *Ruminococcus flavefaciens* strain 17 (Bayer, 2008).

A covalently-linked cellulosome due to the scaffoldin ScaE (Rincon et al., 2005), characterizes *R. flavefaciens* through a sortase-mediated mechanism (Schneewind et al., 1995). The scaffoldin ScaB made by 7 cohesins and is connected to ScaE due to a dockerin in its C-terminus end. ScaA is linked to ScaB cohesin due to a C-terminus dockerin providing a scaffold for a large amount of dockerin-containing enzyme subunits characterized by several cellulolytic activities (Bayer et al., 2008).

Interestingly, ScaA and ScaB lack a cellulose binding domain for the interaction with the substrates, but another protein, CttA, related to ScaE, features this specific domain and would be involved in the cell-substrate adhesion (Fig. 5) (Rincon et al., 2007).

Due to the high specificity of cellulolytic enzymes, *Ruminococcus* can hydrolyze several types of cellulose. Several enzymes have been reported, exo-

beta-1,4-glucanase, endo-beta-1,4-glucanase and cellulodextrinases and the main end products of its cellulolytic activity are cellotriose and cellobiose and a small amount of glucose (Helaszek et al. 1991).

The main volatile fatty acids that can be useful for the ruminant nutrition and metabolism are succinic acid, acetic acid and formic acid.

Relevance of cellulose for bacterial growth, can be evaluated considering the high level of complexity and specificity of cellulosome. Moreover, the amount of *Ruminococcus*-related cells identified in the rumen of several animal, underlines the close relation between host and symbiont and the relevance of this interaction.

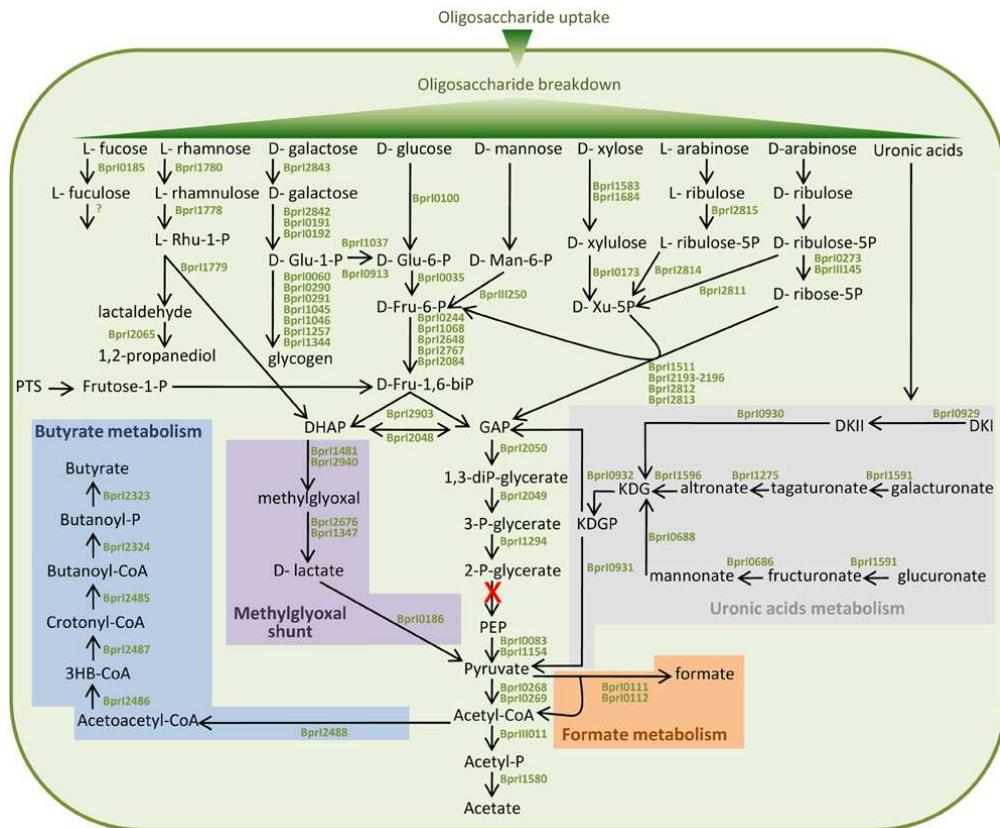
### **3.2.3 *Butyrivibrio* sp.**

*Butyrivibrio* was isolated from the digestive tract of several unrelated organisms and identified for the first time in 1956 as a gram negative curved rod-shaped bacteria motile due to one or more polar to subpolar flagella (Bryant et al., 1956). Morphological evidences revealed that due to the thin cell wall *Butyrivibrio* can be considered a “false Gram negative” (Cheng et al., 1977) but the presence of teichoic acid and the ultrastructure confirms that this bacterium is a Gram positive. Currently, the *Butyrivibrio* genus is belonging to phylum Firmicutes, class Clostridia and counts 4 distinct species ([ncbi.nlm.nih.gov/Taxonomy/](http://ncbi.nlm.nih.gov/Taxonomy/), online source).

Several polysaccharide-degradation encoding genes were annotated in *Butyrivibrio proteoclasticus* B316 genome. An higher specialized glycobiome for cellulose utilization was identified for *B. proteoclasticus* if compared with *F. succinogenes* and *R. flavefaciens* annotation results (Brulc et al., 2009, Miller et al. 2009).

*B. proteoclasticus* does not produce a cellulosome and several cell-associated proteins, are anchored to the peptidoglycan due to a sortase-mediated bridge. Plant polysaccharides like inulin, pectin, starch and xylan may be hydrolyzed by this species (Attwood et al., 1996).

Several ways involved in energy production and metabolism were identified due to the genome annotation (Kelly et al., 2010).



**Fig. 6.** Metabolic reconstruction of the carbohydrate metabolism in *B. proteoclasticus* based on genome annotation. Butyrate and formate represent the main fatty acids due to the methylglyoxal shunt. (Kelly, 2010).

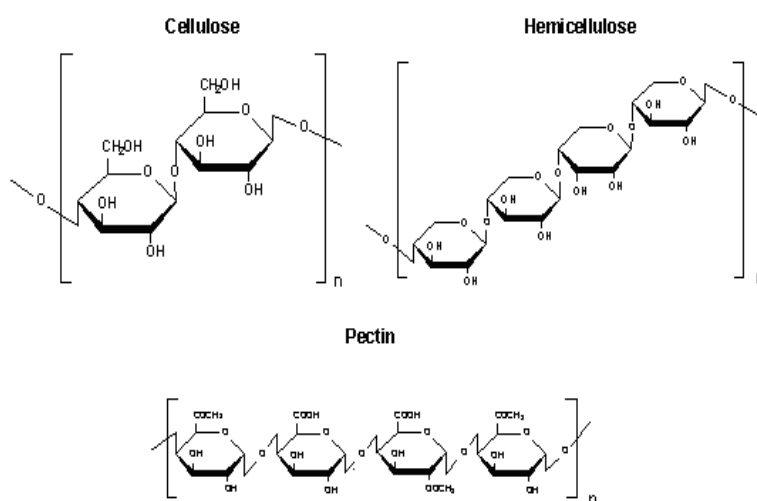
Oligosaccharides breakdown products may enter into different pathways. The lack of an identifiable enolase (the enzyme for the conversion of 2-phosphoglycerate to phosphoenolpyruvate) in the Embden-Meyerhof pathway was observed (Fig. 6). Pyruvate production can occur due to methylglyoxal shunt and specific methylglyoxal synthases catalyze the methylglyoxal production from dihydroxyacetone phosphate. Butyrate and formate represent the main fermentation end products (Attwood et al., 1996). Due to the substrate layers removal ability, *B. proteoclasticus* and other bacterial species like *Prevotella*, were proposed as 'supporting degraders' for the primary degraders *Fibrobacter succinogenes* and *Ruminococcus*.

### 3.3 The digestion mechanism

#### 3.3.1 Cellulose structure and cellulolytic activity

Photosynthetic organisms convert carbon dioxide into carbohydrate due to the photosynthesis. A huge amount of energy is sequestered in the polysaccharide network represented by the plant cell walls due to this pathway.

Several layers of polysaccharidic chains of cellulose cross-linked by hemicellulose constitute the plant cell wall (Lodish et al., 2000) (Fig. 7).



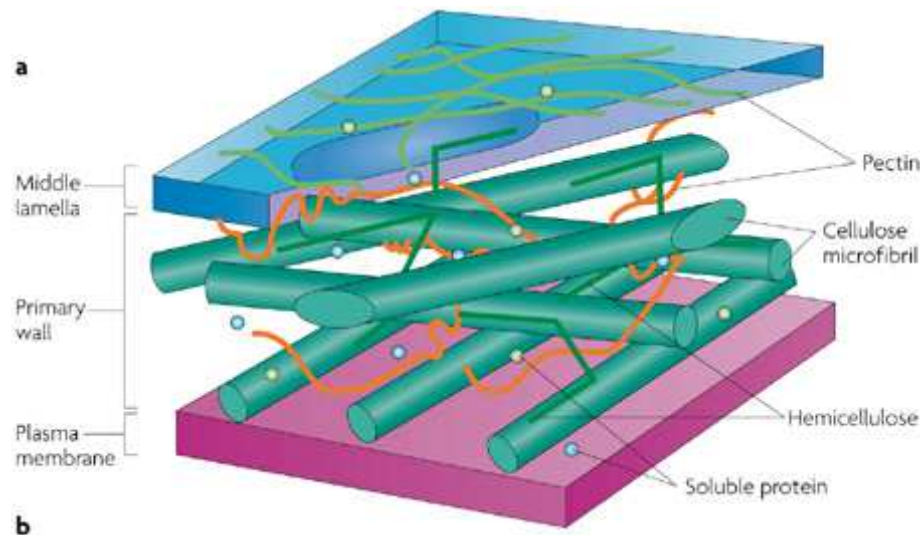
**Fig. 7.** Chemical structures of cellulose, hemicellulose and pectin. (Adapted from <http://www.sigmaaldrich.com>).

Cellulose consists in unbranched and unsubstituted (1,4)-beta-D-glucan chains which can be disposed in a microfibrillar complex due to a huge amount of noncovalent bond and hydrophobic interactions. Other non-cellulosic polysaccharides like pectins or hemicellulose like xyloglucans and heteroxylans (Fig. 7) sustain the plant cell wall and type and chemical features are different between tissues and types of plant (Burton et al., 2010).

Linear chains originated by cellulose polymerization, provide structural support for the cell wall (Pauly et al., 2008) while hemicelluloses increase the mechanical strength due to hydrogen bond linkage with the cellulose matrix (Evans,

Cambridge University Press).

Cellulose, represents a polymer recalcitrant to an efficient deconstruction to simple sugar monomers due to an extremely compact and time-stable structure (Fig. 8).



**Fig. 8.** Structure of the plant cell wall. The cell wall is constituted by cellulose microfibrils, hemicellulose, pectin, lignin and soluble proteins. (Sticklen, 2008).

This lignified and resistant form of plant material, can be hydrolyzed by the consortium of bacteria, fungi and protozoan that characterizes the rumen. Only lignin can not be used in the energy uptake process.

A prolonged chewing time of ingested material represents the first feeding process step. About 25.000 chews for more than 8 hours are required by ruminants (Hume et al., 1980, Mackie, 2002) .

Rumination has several effects: on one hand, the whole properties of forage are maintained by longitudinally shearing and tearing. On the other hand, surface of plant material available for microorganisms increases due to an higher amount of food particles stored in the ruminal cavity with respect than the entire fibers (Weimer et al., 2009)..

Food remains for several hours in the rumen, up to 72 hours for cows underlining the complexity of cellulose digestion despite the complex specialized community (Russell et al., 1992)..

Gut capacity and fermentation mass are isometrically related to the body size: metabolic rate decreases increasing the animal size while the ratio with gut capacity remain unaltered. Turnover rate of the gut content of a larger herbivore is slower respect to a smaller animal and food has a longer residence time for fermentation and digestion of recalcitrant plant material (Mackie et al., 2002).

Cellulolysis requires a close adhesion between microorganisms and substrate. Only a small amount of cellulolytic enzymes have been detected as free molecules by *in vitro* experiments, underlining the inability to digest fibers due to a long-distance process.

A close interaction between bacterial cells and substrates is supported by the presence of cellulosomes and other surface proteins characterized by a cellulose binding module.

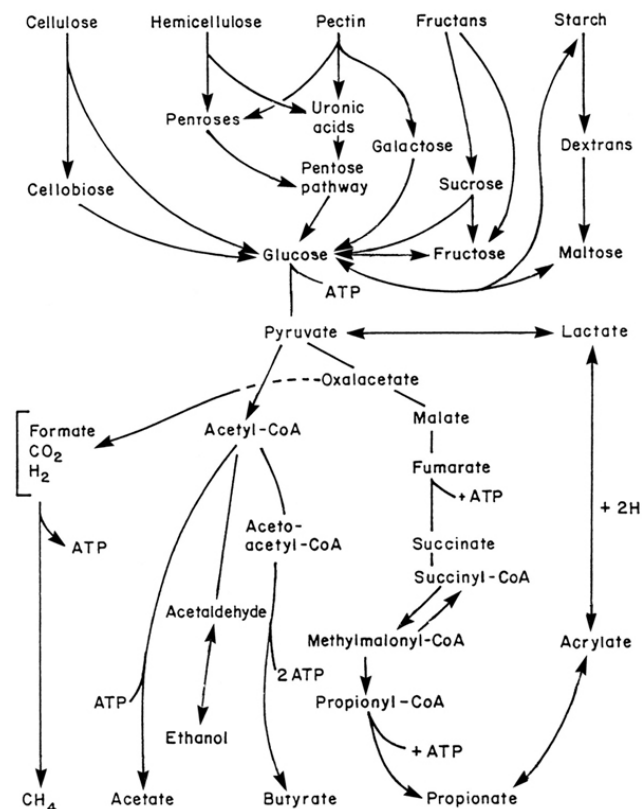
An hypothetical mechanism for cellulolysis was suggested by genomics and *in vitro* studies. *Fibrobacter* seems to be equipped with genes for the complete cellulose degradation, from a xylanolytic activity for the removal of hemicellulose and access to the below cellulose layers. *Ruminococcus species* lack in this functions but a cooperative interaction seems to involve other bacterial species like *Butyrivibrio* in the early stages of fermentation (Koike et al., 2009)..

Several specific cellulases characterized all cellulolytic bacteria: endoglucanases (or 1,4-beta-D-glucan-4-glucanohydrolases) randomly cut internal amorphous sites in the polysaccharide chain generating several oligosaccharides with different lengths and new ends. Exoglucanases, including cellodextrinases (1,4-beta-D-glucan glucanohydrolases) and cellobiohydrolases (1,4-beta-D-glucan cellobiohydrolases) can hydrolyze cellulose starting from the reducing or non reducing end and continuing lysis walking along the carbohydrate chains. Beta glucosidases (beta-glucoside hydrolases) hydrolyze soluble cellodextrins and cellobiose as end products of the exo and endoglucosidases activity, forming glucose (Kelly et al., 2010; Suen et al., 2011).

The current model that explains fiber digestion, provides the cooperative interaction between different cellulases activity (Din et al., 1994). Moreover, an increasing fermentation was observed in heterogeneous communities, suggesting a close relationship between cellulolytic and non-cellulolytic bacteria.

The entire spectrum of plant cell wall polysaccharides (cellulose, hemicellulose, pectin, starch and fructans) could be hydrolyzed and fermented considering the synergism between cellulase and hemicellulase (Fig. 9).

Depending on food type, different amount of volatile fatty acids, CO<sub>2</sub> and methane can be produced by fiber hydrolysis. Volatile fatty acids, represent the main source of energy for the ruminants estimated between 50% and 70% of caloric intake (Sutton et al., 1985; Bergman et al., 1990).



**Fig. 9.** Charbohydrates transformation during the ruminal microflora activity. (Van Soest, 1994).

After production, they are rapidly absorbed by the rumen. Fatty acids transition become without the need of any active transport and occurs due to free energy neutralization determined by the pH gradient between rumen and blood. (Van Soest 1994). A small amount of molecules can be directly converted by the ruminal epithelium but many of these come through the bloodstream into specific organs and enter metabolic pathways .

High levels of propionate in bloodstream, indicate that this molecule is not

metabolized by the ruminal epithelium (Kristensen et al, 2000, Kristensen et al., 2004). In the liver, propionate is converted in oxaloacetate and phosphoenolpyruvate, representing one of the main molecules involved in glucose synthesis. Acetate represents the principal substrate for lipogenesis and fat production and in particular it is involved in the milk fat biosynthesis while butyrate is quickly converted to ketone bodies (Gross et al., 1990; Reynolds, 2002).

### **3.3.2 Protein hydrolysis and digestion**

If cellulolysis represent the basis for ruminants energy production, the digestive tract-associated microbiota is involved also in a modulating nitrogen metabolism, protein assimilation and production. It was estimated that on the basis of food type (Ramos et al., 2009) microbial activity contributes from 40% to 90% of proteins that pass through the rumen to the small intestine (Leng et al., 1984).

Ruminal protozoa play a predominant role in protein degradation (Jouany, 1996) and are partially responsible in modulating the bacterial activity (Koenig et al., 2000).

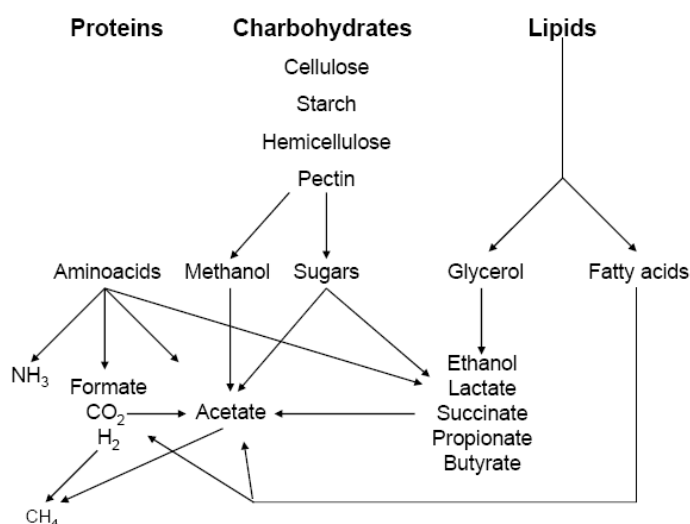
A production of  $\text{NH}_3$  is the consequence of bacterial protein degradation in the rumen. This molecule can be absorbed by the ruminal epithelia and converted to urea (Reynolds, 1995) or involved in aminoacid synthesis (Parker et al., 1995). An amount of produced  $\text{NH}_3$  is absorbed by bacteria themselves (Nolan et al., 2005; Firkins et al., 2007; Reynold et al., 2008).

A different capability in simple or complex proteins hydrolysis, is closely related to the different proteolytic activities that characterize cellulolytics and amylolytics bacteria, (Russell et al., 1992; Wallace et al., 1997).

Once integrated in the cytoplasm, peptides can be hydrolysed producing aminoacids used for microbial protein synthesis, fermented to volatile fatty acids or deaminated and excreted from the cell as ammonia (Tamminga, 1979).

A close relationship was found between growth rate and microbial population requirements of ammonia, peptides, aminoacids as main nitrogen source (Russell et al., 1992).

Microbial biomass is carried from rumen to abomasum and to small intestine during the latest phases of rumination. In this portion, mammalian common digestive and absorptive processes take place and microbes are digested and used by the animals. Hydrolysis of microbial proteins as food components and obtained peptides, represent a fundamental source of nutrition for the animals.



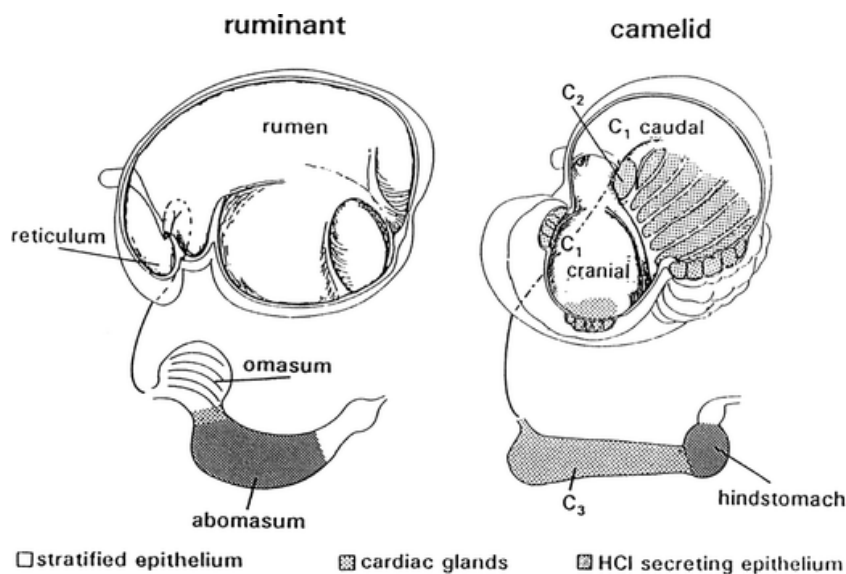
**Fig. 10.** Chemical transformations of proteins, carbohydrates and lipids occurring due to the microbial activity. (Adapted from <http://www.fao.org>).

### 3.4 The dromedary rumen

The digestive capability of camelids in which microorganisms play a fundamental role, represents an interesting biological system which remains largely unexplored. The xerophytic food sources represented by desert plants warrant a possible wealth of microbial-associated enzymatic properties that makes the study of these systems of primary importance for biotechnological purposes.

The harsh environment in which camelids can live, determined the development of several adaptive strategies for the survival in conditions of limited water and food resources.

Camels are defined as *pseudo-ruminants*, characterized by a three-compartmentalized foregut constituted by rumen, reticulum and gastric secreting abomasum instead of four compartments of other ruminants in which the omasum represents an additional chamber (Von Engelhardt et al., 2007) (Fig. 11).



**Fig. 11.** Comparison of the forestomach of ruminants and camelids. C1, C2 and C3 indicate the three compartments that characterize the foreguts of camelids.

The taxonomy of Old World camels has been defined since the first classification of 1758 in which Linnaeus defined *Camelus bactrianus* the two-humped camel and *Camelus dromedarius* the one-humped camel.

Dromedary is diffused in much of the Northern Africa and in regions of the Middle East, occupying environments which are hostile for other organisms (Mares, 1999). Due to the generalist digestive system, camels can survive feeding on salt-tolerant plants and other poorly digestible feed such as shrubs and trees (Kayouli et al., 1993). Digestive capabilities seem to be related to the longer retention and mixing of solid digesta in the forestomach (Lechner-Doll et al., 1991) and to the higher cellulolytic activity of the associated microbiota (Jouani et al., 1995).

A limited number of studies addressed the identification of microorganisms that populate the dromedary rumen and the definition of a digestive pathway is currently under investigation (Buraoui et al., 2006). Two recent works underline that a higher number of Firmicutes and Bacteroidetes are representative of such microflora since the early stages of the dromedary neonatal development (Li et al., 2011; Samsudin et al., 2011). Other studies are required to understand the relationship between these microorganisms and clarify their involvement and contribution to the adaptive capacity of dromedary to survive in adverse

environments and extreme conditions.

### **3.5 Aim of the research**

Rumen ecosystem, represents a natural bioreactor in which hundreds of species cohabit and cooperate in a co-metabolic situation. At present, some of these, are largely studied and characterized by the recent high-throughput sequencing data, revealing that a huge amount of species remain genetically not characterized.

Within the domestic animals that represent the race in productions, several wild ruminants exist and can survive in extreme natural environments in terms of water availability, roughness and fiber content of the food and low carbon/nitrogen ratio.

The specific microbiota, must be selectively adapted to an equally hard and intensive task of conversion in order to satisfy the vital necessities of the animal.

A collaboration with the University of Constantine, Algeria, has underlined the strong degradative rate of bacterial symbionts of Tilopod ruminants reared in semi-wild conditions and the possibility to identify and characterize them (Haddi et al., 2000; Haddi et al., 2003).

Considering the relationships between rumen associated bacteria and adaptive capacity, ruminal microbiota of dromedary was investigated with attention to a possible biotechnological application for the treatment of pectic residues and fibrous rough plans for livestock food production.

In Italy, over 17 million tons of plant biomasses derived from agricultural residues, woody material and agro-industrial wastes are available each year are recycled for energy production (electrical, biogas, bioethanol and biodisel) (Salmaso et al., 2006) The energetic worldwide needs are flanked by the needs of alimentary nature: for several culture, the animal meat represent a fundamental constituent of diet and livestock production represent an important contribution to define a national economic potential. Finding novel alimentary sources at low cost, could represents an interesting aim of the research for which an economical value may be predicted.

Discarded plant biomass offers a good source of material for livestock feed production but, despite the high nutritional value and high equivalence in forage

units, can not be easily digested by bovine and other livestock. The high digestive capability of dromedary which allows the assimilation of fiber-rich vegetables, shrubs and trees, is the consequence of rumen associated microflora. Deciphering the complex interaction within this consortium may allow the development of an *in vitro* fermentation bioreactor for a pre-treatment of discarded indigestible biomass and conversion in easily assimilable compounds for livestock. Proteins of ruminal bacteria, have an higher biological value with respect to the digested substrates and represent a fundamental source for animals: a modulated pre-digestion process may be used to optimize the protein supply for livestock increasing the efficiency of nitrogen assimilation and allowing a control of post nutritional nitrogen compounds emission into the environment.

For this purpose, 16S rRNA genes from total DNA extracted by the rumen were amplified due to universal primers and sequenced due to a 454-GL FLX upgraded Titanium. To analyze the effect of diet in the modulation of bacterial population, two dromedaries fed by different substrates were considered: an hay-fed animal and an atriplex-fed animal. Atriplex is a xerophytic dicotyledonous plant to the family Amaranthaceae. While hay represent the most commonly used food for livestock, the fibrous atriplex is be a natural food for a wild dromedary determining a possible microbial enrichment for which the species characterized by an higher cellulolytic activity may be the most representative.

A comparison between microorganisms developed during the different digestive processes, can allow to understand the interaction and the possible role for each taxon as well as the identification of species which can be useful for biotechnological application.

## Chapter 4. Materials and Methods

### 4.1 List of Abbreviations and Terms Used

AMPure®	Agencourt® AMPure® DNA purification kit
bp	base pair
BSA	Bovine Serum Albumin
COD	Chemical Oxygen Demand
CTAB	Cetyltrimethylammonium Bromide
dNTPs	Deoxynucleoside Triphosphates
ds	double strand
Eppendorf	Centrifuge tubes supplied by Eppendorf®
EDTA	Ethylene-Diamine Tetraacetic Acid
EtBr	Ethidium Bromide
f.c.	final concentration
LiCl	Lithium Chloride
NaAc	Sodium acetate
NH <sub>4</sub> Ac	Ammonium acetate
SDS	Sodium Dodecyl Sulphate
TAE	Tris-Acetate-EDTA
TE	Tris-EDTA
Tris	2-amino-2-hydroxymethyl-1,3-propanediol

## 4.2 General Solutions and Preparations

### DNA extraction

<b>CTAB 10% v/v</b>	(100 ml)
NaCl	4.1 g
CTAB	10 g
Water	to 100 ml

### Gel electrophoresis

<b>Tris Acetate</b>	(100 ml)
Tris base	12.1 g
Acetic acid	2.855 g

### **0.5 M EDTA**

EDTA disodic dehydrate	9.3 g
NaOH pills	1.0 g
Adjust pH to 8.0 adding NaOH 10 N.	

### **TAE buffer 50X** (1 L) (running electrophoresis buffer, gel component)

Tris base	242 g
Glacial acetic acid	57.1 ml
EDTA 0.5 M (pH 8.0)	100 ml
mQ water	to 1 L

<b>Agarose Gel 1%</b>	<b>(50 ml)</b>
Agarose	0.5 g
Filtered TAE 1X buffer	50 ml
EtBr 20000X	2.5 µl

<b>Agarose Gel 2%</b>	<b>(65 ml)</b>
Agarose	1.3 g
Filtered TAE 1X buffer	65 ml
EtBr staining after the complete run.	

### **4.3 Sample preparation for electron microscopy**

Electron microscopy was carried out in collaboration with Professor Barbara Baldan, Dept. of Biology of the University of Padua.

#### **4.3.1 Scanning Electron Microscopy**

The specimens for SEM were prepared by fixing with a 2,5% glutaraldehyde in 0,2 cacodylate buffer for 4 hours. After that, samples were rinsed three times in the same buffer at 4°C for 4 hours each and post-fixed in osmium tetroxide for 30 minutes. Three washes in 0.1M cacodylate buffer for 30 minutes were carried out followed by dehydration with a greater series of ethanol and dried at critical point. A treatment with hexamethyldisilazane for 1 hour was performed and samples were dried for 1 hour at room temperature under a fume hood.

Finally, samples were coated with gold in an Edwards S. 150B sputter coater and examined using a Cambridge Stereoscan 250 Scanning electron microscope.

#### **4.3.2 Transmission Electron Microscopy**

Specimens for TEM were prepared by fixing with 3% glutaraldehyde in 0.1 M cacodylate buffer at 4°C for 24 h. Specimens were rinsed three times in the same buffer for 30 minutes each and post-fixed in osmium tetroxide 1% at 4°C for 2 h. Three washes in 0.1 M cacodylate buffer for 30 minutes were carried out, followed by dehydration with a graded series of ethanol and dried at critical point. Specimens were then rinsed three times in propylene oxide for 25 minutes, included in resin and dried at 40°C for 24 h and at 60°C for three days. 0.05 µm sections were prepared in copper nets and contrasted with lead citrate for 30 minutes.

All sections, were viewed using a FEI-Tecnal 12G transmission microscope operated at 75-100 kV.

### **4.3.3 Negative staining**

A standard protocol for the negative staining was applied: carbon-coated 400 mesh copper grids were deposited on a drop of sample suspension for 10 minutes and rinsed twice with water removing the fluid phase using a filter paper wedge. Afterwards, the grids were stained for 10 seconds on a drop of 2% uranyl acetate, dried and analyzed using a Tecnai TEM (FEI Company, Eindhoven, the Netherlands) operating at an acceleration voltage of 120 kV.

## **4.4 Samples analysed**

### **4.4.1 The dromedary rumen**

Ruminal samples were collected by Dr. Haddi Mohammed, a collaborator and professor at the Université Mentouri of Constantine, Algeria. The dromedaries reared in semi-wild state near the desert slopes of the chain of the Sahara Atlas. D1 dromedary was fed by an hay-based diet for 1 week before sampling while D1 was fed by atriplex, a natural source for such organisms.

### **4.4.2 The Anammox samples**

Biomasses were collected directly by the bioreactor of the Eurotec Water Treatment Technologies in Padua. 2 ml of sample were withdrawn and centrifuged for 5 minutes at 10000 rpm to compact the cells. Samples were frozen in liquid nitrogen and stored at -80°C until the DNA extraction.

### **4.4.3 Milan pilot bioreactor**

Biomasses collected from the Eurotec WTT pilot bioreactor were also used by Professor Adani research group (University of Milan) to set up another pilot bioreactor. Using sludges collected from this reactor and from the Eurotec WTT reactor, distinct reactor batches with differences in the nourishing solutions were set up. More specifically, the following inocula were used:

- biomass from the Milan pilot bioreactor (79 days of activity);
- biomass from the Eurotec WTT pilot bioreactor kept refrigerated at 4°C;

- biomass from the Eurotec WTT pilot bioreactor kept frozen at -20°C;
- granules from the Eurotec WTT pilot bioreactor collected during its
- highest anammox activity and kept frozen at -20°C.

Characteristics of the analyzed samples are listed in Tab. 1. Grey-shaded boxes, indicate samples for which a 454-Roche sequencing was carried out.

Sample	Biomass	Notes
A1	biomass from Milan pilot bioreactor	Initial COD/N rate is equal to 1 to selectively grow anammox bacteria instead of denitrifying microorganisms.
A2	biomass from Milan pilot bioreactor	Initial COD/N rate is equal to 1 to selectively grow anammox bacteria instead of denitrifying microorganisms.
A5	biomass from Milan pilot bioreactor	A1 with addition of glycerin to analyze the effect of organic matter on microbial population
A6	biomass from Milan pilot bioreactor	A1 with addition of glycerin to analyze the effect of organic matter on microbial population
B1	/	Blank (no inoculum)
B2	/	Blank (no inoculum)
C1	biomass from Milan pilot bioreactor	Initial COD/N rate is equal to 2 to assess if such conditions would affect negatively the anammox population.
C2	biomass from Milan pilot bioreactor	Initial COD/N rate is equal to 2 to assess if such conditions would affect negatively the anammox population.
D1	biomass from denitrifying plant	Denitrifying bacteria surely present ( to mcompare with A5 - A6)
D2	biomass from denitrifying plant	Denitrifying bacteria surely present (to compare with A5 - A6)
E	Eurotec inoculum kept frozen (-20°C)	
E2	Eurotec inoculum kept frozen (-20°C)	E with addition of glycerin to analyze the effect of organic matter on microbial population
F	Anammox granules	
G	Eurotec inoculum kept refrigerated (4°C)	
amx	Eurotec bioreactor	Samples collected from the Eurotec WTT pilot bioreactor.

**Tab.1.** List of analysed samples and their specifics.

COD is the Chemical Oxygen Demand: it measures the amount of organic compounds in water. A low COD/Nitrogen ratio would favour anammox bacteria (Güven, 2005), while a high COD/N rate would negatively affect them.

Samples A to G came from the Milan research group bioreactors and were collected at two distinct time points, day 25 and day 65 from batch start up. To distinguish sampling times, a .1 and .2 annotation was adopted. For example, A6.1 and A6.2 refers to the same bioreactor, A6, but the samples were collected at the two different time points (Dates: 12/21/2010 - 01/25/2011).

The sample identified by the 'amx' abbreviation comes from the Eurotec WTT anammox pilot bioreactor. Sampling at four different time points was done and these are indicated as 1, 2, 3 and 4 (Dates: 01/22/2009 - 02/03/2009 - 03/03/2009 - 04/07/2009).

Sequenced samples are highlighted in grey.

#### **4.5 Rumen and Anammox samples: total DNA extraction**

Commercial kits specifically improved to ensure maximum DNA yield from environmental samples are widely available. The PowerSoil® DNA Isolation MOBIO Kit for genomic DNA isolation from soil has been used and optimized for these kinds of samples. This kit uses a protocol which separates humic acids, the removal of which is preferable as they negatively affect all subsequent reactions. The kit provides mechanical and chemical cell lysis, using rock beads and standard detergents such as SDS, or enzymes such as lysozyme. Some solutions are used and samples are subsequently centrifuged. Finally the kit supplies spin filters for DNA purification with a silica membrane.

To achieve a higher DNA yield several steps of the protocol were modified. In particular, a proteinase-K treatment for the Eurotec WTT Anammox samples and those from the Milan bioreactors and a phenol:chloroform:isoamylalcohol extraction step, also used for rumen samples, were introduced.

Comparisons between DNA amount and purity have been made to decide the

optimal extraction method.

Sample	Sample preparation
Dromedary D1 and E1	0,3 mg of lyophilized sample were collected in a clean 2 ml eppendorf tube
Eurotec WTT amx samples and Milan samples	1,5 mg of bioreactor were pelleted in a clean 2 ml eppendorf tube after centrifugation at 13.000 rpm.

**Tab. 2.** Amount of collected sample

Following the PowerSoil® DNA Isolation MOBIO Kit user manual, the recommended protocol was tested: DNA was always resuspended in 100 µl of water.

Yield was evaluated loading 5 µl of sample onto a 1% agarose gel, using 1-kb DNA ladder and running it for 30 minutes at 120 V.

Yield and purity were also confirmed due to a Thermo Scientific NanoDrop® ND-1000 UV-Vis spectrophotometer.

#### 4.5.1 Protocol Modifications

Several alternative protocols were tested using lysis solutions of the PowerSoil® DNA Isolation MOBIO Kit. Bead beating was always applied for the granular sample desruption and other experimental procedures were integrated to lysis and cleaning buffer C1, C2 and C3. Moreover, phenol:chloroform:isoamylalcohol (25:24:1) purification and ethanol-salt precipitation, were always preferred to the silica spin filters supplied by the MOBIO kit because of the low recovery of DNA observed with protein- and carbohydrate-rich samples.

Even if genomic DNA was always recovered, different yields and quality could be related to the different chemical steps encountered. Samples extracted following the entire MOBIO kit protocol displayed a lowest DNA yield but the highest DNA purity for all samples.

The optimized protocols for rumen and bioreactors DNA extraction are described below:

#### 4.5.2 Dromedary rumen DNA extraction protocol

- The Resuspension Buffer was added without beads and vortexed for 5 minutes to hydrate sample: then, beads were added to the tube and vortexed for 5 minutes to resuspend the solution.
- C1 solution (containing SDS) was added and the suspension obtained was placed on a vortex for 15 minutes.
- Samples were centrifuged at RT for 10 minutes and solution C2 was added to the collected supernatant. After 5 minutes of incubation at 4°C. samples were again centrifuged for 5 minutes at RT, at 10000 rpm.
- Solution C3 was then added to the supernatant and the suspension was incubated for 5 minutes at 4°C, centrifuged for 5 minutes at RT at 10000 rpm, then the supernatant was recovered and transferred into a new eppendorf tube.
- An isovolume of phenol:chloroform:isoamylalcohol (25:24:1) was then added. Samples were centrifuged for 5 minutes at RT. This step and the ones which follow can be performed in phase lock tubes (QIAGEN MaXtract® High Density). The phenol:chloroform:isoamylalcohol extraction was repeated.
- An isovolume of chloroform was added and the solution was centrifuged.
- The aqueous phase was recovered and precipitated with ethanol and salts. NH<sub>4</sub>Ac to a final concentration (f.c.) of 2M and 2.5 volumes of 100% filtered ethanol were added.
- Samples were left at -20°C overnight and then centrifuged for 1 hour at 4°C and 13000 rpm. The supernatant was removed.
- Pellets were washed with decreasing amounts of 75% ethanol and spun three times.

- After drying the pellets by leaving the tubes opened in the laminar flow hood, they were resuspended in the desired volume of water.

### **4.5.3 Anammox and Milan bioreactors DNA extraction**

While the first steps were the same used for dromedary samples (C1, C2 and C3) a good result was obtained integrating a proteinase-K treatment after the three MOBIO solutions.

- After adding TE buffer to samples, proteinase K (20 mg/ml) was added to reach a final concentration of 100 µg/µl. Samples were incubated at 55°C for 1 hour.
- An isovolume of phenol:chloroform:isoamylalcohol (25:24:1) was then added. Samples were centrifuged. This step and the ones which follow can be performed in phase lock tubes (QIAGEN MaXtract™ High Density). The phenol:chloroform:isoamylalcohol extraction was repeated.
- An isovolume of chloroform was added and the solution was centrifuged.
- The aqueous phase was recovered and precipitated with ethanol and salts. In particular NH<sub>4</sub>Ac to a final concentration (f.c.) of 2M and 2.5 volumes of 100% filtered ethanol were added.
- Samples were left at -20°C overnight and then centrifuged for 1 hour at 4°C. The supernatant was removed.
- Pellets were washed with decreasing amounts of 75% ethanol and spun three times.
- After drying the pellets by leaving the tubes opened in the laminar flow hood, they were resuspended in the desired volume of water.

#### 4.6 Polymerase Chain Reaction (PCR)

The PCR reaction represented one of the central techniques employed in of this work.

Considering the wide number of commercial polymerases, two different enzymes were used for different applications.

For RFLP analysis and semiquantitative PCR a Promega GoTaq® Flexi DNA Polymerase without a proofreading activity was used.

For the sequencing analysis, on the other hand, as an accurate polymerization reaction is needed, because every amplicon will potentially produce a sequence, and consequently a single datum. In this case a Finnzymes Phusion® High-Fidelity DNA Polymerase was preferred. This is a polymerase first isolated from *Pyrococcus furiosus*, fused with a processivity-enhancing domain that increases enzyme fidelity and speed.

The general protocol for a PCR amplification consists of the following steps:

For Promega GoTaq® Flexi DNA Polymerase, see Tab. 3.

Component	Final Concentration	Volume
5X GoTaq® Flexi Buffer	1X	4 µl
MgCl <sub>2</sub> Solution, 25 mM	1.25 mM	1 µl
dNTPs, 10 mM each	0.2 mM	0.4 µl
Forward Primer, 10 µM	0.3 µM	0.6 µl
Reverse Primer, 10 µM	0.3 µM	0.6 µl
GoTaq® DNA Polymerase (5u/µl)	0.5 u	0.1 µl
Template DNA	5 - 25 ng/µl	X µl
Nuclease-free water		up to volume
Total		20 µl

**Tab. 3** Reagents for a PCR with Promega GoTaq® Flexi DNA Polymerase.

Thermal cycler program:

- initial denaturation: 2 minutes at 95°C;
- denaturation: 20 seconds at 95°C;
- annealing: 60 seconds/kb at 55°-65°C;
- extension: from 25 seconds to 1 minute and 30 seconds at 72°C;
- final extension: 6 minutes at 72°C.

27 cycles of denaturation, annealing, extension are completed.

Annealing temperature and time were carefully evaluated on the basis of primer sequence and length. Extension time was also optimized considering the amplicon length. Results were always visualized by agarose gel electrophoresis.

For Phusion® High-Fidelity DNA Polymerase see Tab. 4.

<b>Component</b>	<b>Final Concentration</b>	<b>Volume</b>
Phusion 5X HF Buffer	1X	4 µl
dNTPs, 10 mM each	0.2 mM	0.4 µl
Forward Primer, 10 µM	0.3 µM	0.6 µl
Reverse Primer, 10 µM	0.3 µM	0.6 µl
GoTaq® DNA Polymerase (5u/µl)	0.5 u	0.1 µl
Template DNA	5 - 25 ng/µl	X µl
Nuclease-free water		up to volume
Total		20 µl

Table 4 Reagents for a PCR with Phusion® High-Fidelity DNA Polymerase.

Reagents were mixed paying attention to the hyper-salinity buffer, which may produce bubbles if shaken too much. The thermal cycler program was set as follows:

- initial denaturation: 2 minutes at 98°C;
- denaturation: 20 seconds at 98°C;
- annealing: 45 seconds at 61.5°C;
- extension: from 15 seconds to 1 minute and 30 seconds at 72°C;
- final extension: 6 minutes at 72°C.

25 cycles of denaturation, annealing, extension.

In comparison with the PCR carried out with GoTaq® DNA Polymerase, all cycle temperatures were 3°C higher, due to the hyper-salinity buffer.

#### **4.6.1 Semiquantitative PCR**

Assuming that the number of DNA copies doubles at every PCR cycle, the more amplifiable DNA template is present in the sample, the lower the number of cycles sufficient to have a visible PCR product.

In semi-quantitative PCR, the same PCR mix solution was aliquoted in PCR tubes. They were removed from the thermal cycler at precise time-points and the reaction was stopped by placing the PCR tubes at -20°C. If different samples were compared, their DNA concentrations was normalized to the same value. Under these circumstances, every PCR starts from the same amount of template and comparison between different samples can be made.

Several replicates were set up and, every 3 cycles one tubes was removed.

In this technique if different primers are used, amplicons of the same length have to be produced to obtain comparable results.

#### **4.6.2 PCR primers for ribosomal genes**

Different combinations of primers were tested, according to the type of information to be pursued.

Moreover, new pairs of universal primers were evaluated (in collaboration with Dr. Ivano Zara) using a pair as 'fusion primers' for the 454-Roche pyrosequencer.

### 4.6.3 Universal primers definition

A dataset containing up to 150000 sequences predicted as 16S rRNA was downloaded from the RDP site Release-8 in september 2009 (<http://rdp.cme.msu.edu>) checking all the options for the higher number of queries.

The database was purified using *Escherichia coli* 16S rRNA as probe and removing all the entries which were not recognized as 16S genes.

Sequences were aligned using nucleotide BLAST and *Escherichia coli* 16S ribosomal gene as reference and the conserved regions were analyzed using a c++ *ad hoc* program.

Any potential pair of primer was considered when the distance between the forward and reverse primer was at least 400 bases. In addition, taking into account the different information related to the 16S rRNA genes hypervariable regions, V3 V4 and V5 primers were favored.

Selected primers were named F357, F527, R790, R1064 and their sequences are displayed below. These primers were used for 454-Roche sequencing and for other PCR-based techniques.

Primer features are listed in Table 5.

**Universal primers:** these match conserved regions of the bacterial 16S ribosomal gene. They are useful to investigate the whole bacterial community. The primer sets that have been tested are listed in Tab. 5.

Name	Sequence	Specificity	Reference
F27	AGAGTTTGATCCTGGCTCAG	Bacteria	Hugenholtz and Gobel, 2001
F357	TACGGGAGGCAGCAG TACGGGAGGCTGCAG TACGGGAGGCCGCAG	Bacteria	CRIBI
F527	GCTGCCAGCAGCCG CGTGCCAGCAGCTG TGTGCCAGCAGTCG	Bacteria	Hugenholtz and Gobel, 2001
F818	ATGGGCGACTMRGTAGAGGG GTTT	Anammox bacteria	Tsushima, 2007
Pla46	GGATTAGGCATGCAAGTC	Planctomycetes	Neef, 1998
R790	GTGGACTACCAGGGTATCT CAGGACTACCGGGTATCT TAGGACTACCCGGGTATCT	Bacteria	CRIBI

Name	Sequence	Specificity	Reference
R1064	CACGACACGAGCTGAC CACGGCACGAGCTGAC CACAACACGACGAGCTGAC	Bacteria	Hugenholtz and Gobel, 2001
R1391	GACGGGCRGTGWGTRCA	Bacteria	Hugenholtz and Gobel, 2001

**Tab. 5** List of primer features.

Primers	16S hypervariable region amplified	Amplicon length	Purpose
F357 - R790	V3, V4	? 400 bp	To assess the general biodiversity of the microbiological community present in the samples. The amplicon library obtained via PCR was sequenced via 454-pyrosequencing.
F357 - R1391	V3, V4, V5, V6, V7, V8	? 1000 bp	To investigate the time-dependent development of the microbial community and to evidence differences in the microbial composition of distinct bioreactors. The amplicons obtained are restricted with different restriction enzymes for RFLP analysis.
F27 - R1391	V1, V2, V3, V4, V5, V6, V7, V8	? 1300 bp	To investigate the time-dependent development of the microbial community and to evidence differences in the microbial composition of distinct bioreactors. The amplicons obtained are restricted with different restriction enzymes for RFLP analysis.

**Tab. 6.1** List of universal primer sets tested.

Primers	16S hypervariable region amplified	Amplicon length	Purpose
F527 - R790	V4	? 200 bp	To assess the overall amount of bacteria in the microbiological community. A semi-quantitative PCR was performed with this couple of primers.

**Tab. 6.2** List of universal primer sets tested.

- **Planctomycetes-specific primers:** match conserved regions of the Planctomycetales 16S ribosomal gene. They are useful to investigate the anammox community.

The primer sets that have been tested are listed in Tab. 7.

<b>Primers</b>	<b>16S hypervariable region amplified</b>	<b>Amplicon length</b>	<b>Purpose</b>
Pla46 - R1391	V1, V2, V3, V4, V5, V6, V7, V8	? 1300 bp	To assess the general amount of Planctomycetes in the bioreactors. A semi-quantitative PCR was performed with this couple of primers.
Pla46 - R1064	V1, V2, V3, V4, V5, V6	? 1000 bp	To investigate the time-dependent development of the Planctomycetales community and to evidence differences in the microbial composition of distinct bioreactors. The amplicons obtained are restricted with different restriction enzymes for RFLP analysis.
F818 - R1064	V5, V6	? 200 bp	To assess the general amount of Planctomycetes in the bioreactors. A semi-quantitative PCR was performed with this couple of primers.

**Tab. 7** List of Planctomycetes- or anammox- specific primer sets tested.

#### **4.6.4 PCR-products purification**

AMPure® is a commercial kit specifically designed for DNA or RNA purification, which is necessary when buffer solutions have to be changed due to their incompatibility with downstream reactions, or when nucleotides or oligos (e.g. primers) are present which may interfere with subsequent analysis.

It consists of magnetic beads that selectively bind nucleic acids longer than 100 bp.

The general steps are listed below:

- AMPure beads were added in proportion to sample volume (1.8 µl of beads per µl of sample) and mixed.
- Beads and DNA were incubated at room temperature for 5 minutes to allow nucleic acids to bind.

-The suspension was then placed on a magnetic support for at least 2 minutes to allow separation of beads and DNA from the remaining solution. The solution was removed.

-Two washes with 200  $\mu$ l of 70% ethanol were carried out.

-Beads were left drying for 5 minutes to allow ethanol evaporation and then eluted with nuclease-free water.

#### 4.6.5 PCR product lyophilization

Before sending them to BMR genomics (rumen D1 and E1) and to the Ramaciotti Center (Anammox and Milan samples) for 454-sequencing, samples were frozen at  $-80^{\circ}\text{C}$  and lyophilized

#### 4.7 RFLP

Restriction Fragment Length Polymorphism is a technique used to compare samples from different bioreactors and collected at different time-points. This method can provide important information concerning the general microbial community composition or its development in time.

It consists of a digestion of 16S rRNA gene amplicons with different restriction enzymes and their subsequent visualization in a gel electrophoresis. The result is an ensemble of DNA fragments of different sizes that generate a specific pattern for that sample or time point. When the bacteria community changes in its composition, the electrophoresis pattern will also change, due to the alteration in relative amplicon and fragment amounts.

The restriction enzymes used are listed in Table 3.6.

Enzyme	Recognized sequence	Overhang	Incubation	Inactivation	Stock concentration	BSA
<i>Sau3A</i>	GATC	5'	$37^{\circ}\text{C}$	$65^{\circ}\text{C}$ per 20 minutes	4000 u/ $\mu$ l	yes
<i>HaeIII</i>	GGCC	Blunt	$37^{\circ}\text{C}$	$80^{\circ}\text{C}$ per 20 minutes	10000 u/ $\mu$ l	no

<i>RsaI</i>	GTAC	Blunt	37°C	65°C per 20 minutes	10000 u/μl	no
<i>HinfI</i>	GANTC	5'	37°C	65°C per 20 minutes	10000 u/μl	no
<i>HindIII</i>	AAGCTT	5'	37°C	80°C per 20 minutes	20000 u/μl	yes
<i>HgaI</i>	GACGC	5'	37°C	65°C per 20 minutes	2000 u/μl	no

**Table 3.6** List of restriction enzymes features.

Reagent solutions for each enzyme are listed in Table 3.7.

Component	Final concentration	Volume
Buffer 10X	1X	2 μl
Enzyme	1 u per μg of DNA to be digested in 1 hour	To be determined depending on PCR product concentration
BSA 100X	1X	0.2 μl
PCR product	To be determined depending on PCR result	To be determined depending on PCR result
Nuclease-free water		up to volume
Total		20 μl

**Tab. 8** List of reagents for a restriction digestion reaction.

After digestion has occurred, a 2% agarose gel is made to visualize the results. The electrophoresis must be performed at 4°C, with a voltage of 50 mV for 4 - 5 hours, to allow DNA fragments to separate clearly.

## 4.8 Bioinformatic tools

### 4.8.1 BLAST

Blast (Basic Local Alignment Search Tool) finds regions of local similarity between sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance of matches. BLAST

can be used to infer functional and evolutionary relationships between sequences as well as to help identify members of gene families..

Blast can work using different matrices: the alignment proceeds via scanning windows and when a similarity region is found, the program tries to expand it laterally. Through a local alignment, Blast finds the subject sequence (or the sequences) which more closely match the query sequence. The alignment must be performed against a database of sequences, chosen based of the purpose of the analysis.

Blast can manage multiFASTA format of sequences as input and always needs a reference database. The default algorithm parameters were modified as follows:

- e-value: expect value; this setting specifies the statistical significance threshold for reporting matches against database sequences. For all analysis an e-value of 0.05 was used.
- Number of best hits to display.
- Release 2.2.22 was used.

#### **4.8.2 ClustalW**

ClustalW is a general multiple sequence alignment program for DNA or proteins. It aligns three or more sequences together in a computationally efficient manner, allowing for their entirety. Multiple sequence alignment is a significant development in phylogenetic analysis, which aims at modeling the substitutions that have occurred over evolution deriving the evolutionary relationships between sequences. The program does not require any reference database.

Release 2.1 was used

#### **4.8.3 CdHIT**

This program performs cluster analysis on multifasta files. Once a cluster threshold is chosen, sequences are grouped in relation to their similarity. The program generates two different files: one with sequences grouped by cluster, one with a multifasta of only the most representative sequences for each cluster.

CdHIT is useful to speed up analysis on huge datasets, because once sequences are grouped, subsequent analysis can be made on the most representative ones only. Release 4.5.5 was used.

#### **4.8.4 Perl**

Perl is a general-purpose programming language helpful for text manipulation as a BLAST or ClustalW output.

It allows to manage very large amount of data as in the case of 454-reads in a relatively easy and quick way, automating different actions, for instance storing Blast results in MySQL tables .

Perl was used to write scripts for the following purposes:

- to perform different operations on sequences (for example convert .xml format of Blast results obtained with Camera to fasta format, trim and annotation sequences, graph creation etc...);
- to automatically load data in MySQL databases (for example Blast results);
- to automatically extract information from MySQL databases (for example `taxa_count.pl`: a script which extracts taxonomical classification for each read and counts the number of reads which can be traced to a determined taxa; `database_comparison.pl`: a script which counts the number of sequence matches for the same species against different databases; `ribomatrix.pl`: `J24_dizzy.pl`: script that creates a graph showing minimum, maximum and average similarity value for all the organisms or query within a specified taxa or taxonomic level; `virtualprinting.pl`: script that produces a matrix displaying a list of the fragment obtained by a 'virtual' digestion of a group of sequences etc.).

#### **4.8.5 MySQL**

MySQL is an open source multithreaded, multi-user SQL database management system (DBMS). It enables to store, modify and extract information from a database and can be interfaced with a Perl script.

Information is stored in Tables made of different fields called records that can store distinct types of data, from numbers to text. A single MySQL database can contain many interconnected and interacting tables at once and store thousands of individual records.

In this study, MySQL was used to store NCBI Taxonomy database, Blast results and to compare them when obtained with different databases.

Release 5.6 was used.

#### **4.8.6 Online tools: Camera, Galaxy and RDP**

Camera and Galaxy are online tools that allow users to perform a set of analysis on next-generation sequencer reads and align sequences against various reference databases.

Camera (Community Cyberinfrastructure for Advanced Microbial Ecology Research & Analysis) was developed by the California Institute for Telecommunications and Information Technology (<https://portal.camera.calit2.net>). This tool can be used to annotate and cluster metagenomic data, launch Blast analysis, perform statistical analysis on data...

Galaxy (<http://main.g2.bx.psu.edu/>) can be used to trim 454-reads on the basis of their quality values and length, to perform simple text manipulation on FASTA files, to convert files into different formats, to align sequences, to subject 454-reads to megablast, to perform statistical analysis, to fetch taxonomic representation, to draw phylogenetic trees...

These tools are especially useful in that they allow to treat metagenomic data to various analyses difficult to carry out because of the large dataset dimension.

The Ribosomal Database Project (<http://rdp.cme.msu.edu>), provides ribosome-related data a

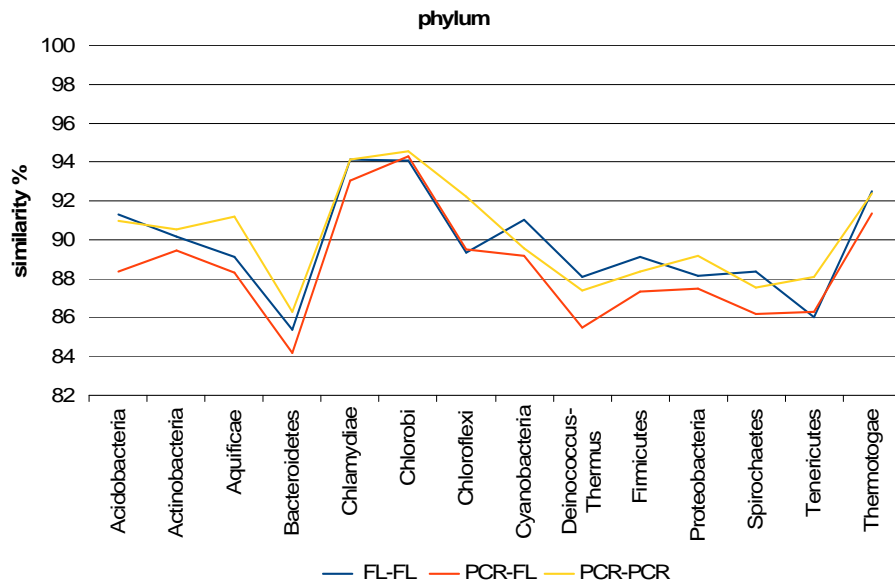
## **Chapter 5. 16S rRNA genes analysis**

### **5.1 Comparison between 16S rRNA genes from sequenced bacteria**

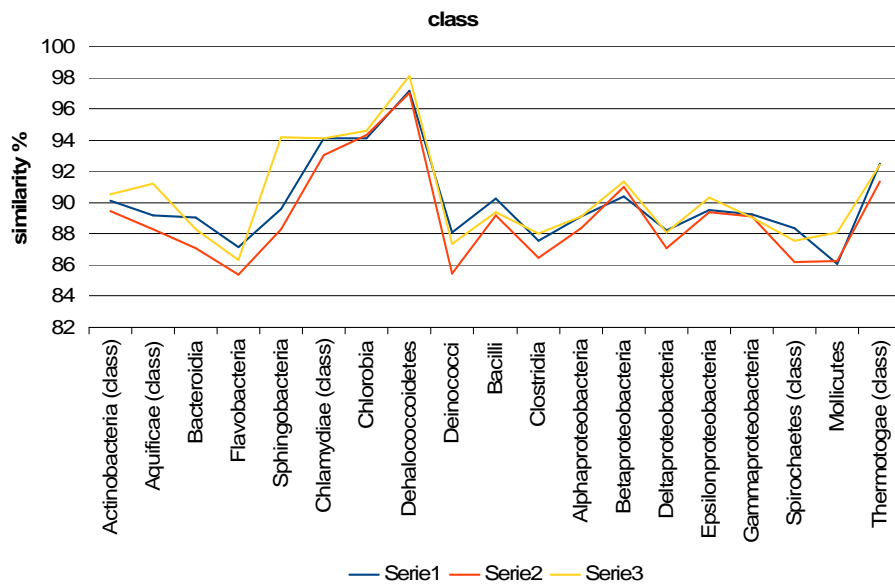
Thresholds to define higher taxonomical levels are presently discussed and represent an important achievement if a 16S rRNA genes analysis is used to envisage a microbial community. To determine if a numerical assessment could be evaluated for phylum, class, order, family and genus annotation, 1045 bacterial genomes have been envisaged using nucleotide BLAST as an ideal data set for such research.

Classification of these species is based not only on the partial information derived by ribosomal sequencing: phylogenetic inferences, biochemical data, serological information and other features which could be considered expression of a polyphasic approach were used and published. But a bias related to the unbalance of sequenced bacteria must be considered: 245 Gammaproteobacteria representatives versus the two soles species related to Planctomycetes and Synergistetes phyla.

To obtain a numerical value for which the resulted 454 sequences of the anammox and ruminal 454 sequences research could be compared, three different BLAST were made giving a similar results for each taxonomic level considered (Fig. 1.1, 1.2). Firstly, full-length 16S rRNA genes were compared. To analyze the information obtained if only hypervariable regions V3 and V4 were considered, sequences comprises between F357 and R790 primers (used for the amplicon library construction) were used to carried out a further analysis. In the third analysis, virtually-amplified sub-sequences were used as query and subject.



**Fig. 1.1** Comparison between the similarity values calculated for different phyla. FL-FL: full length 16S rRNA as query and subject; PCR-FL: sub-sequences obtained by a virtual PCR on 16S rRNA genes as query and full length genes as subject; PCR-PCR: sub-sequence used as query and subject.

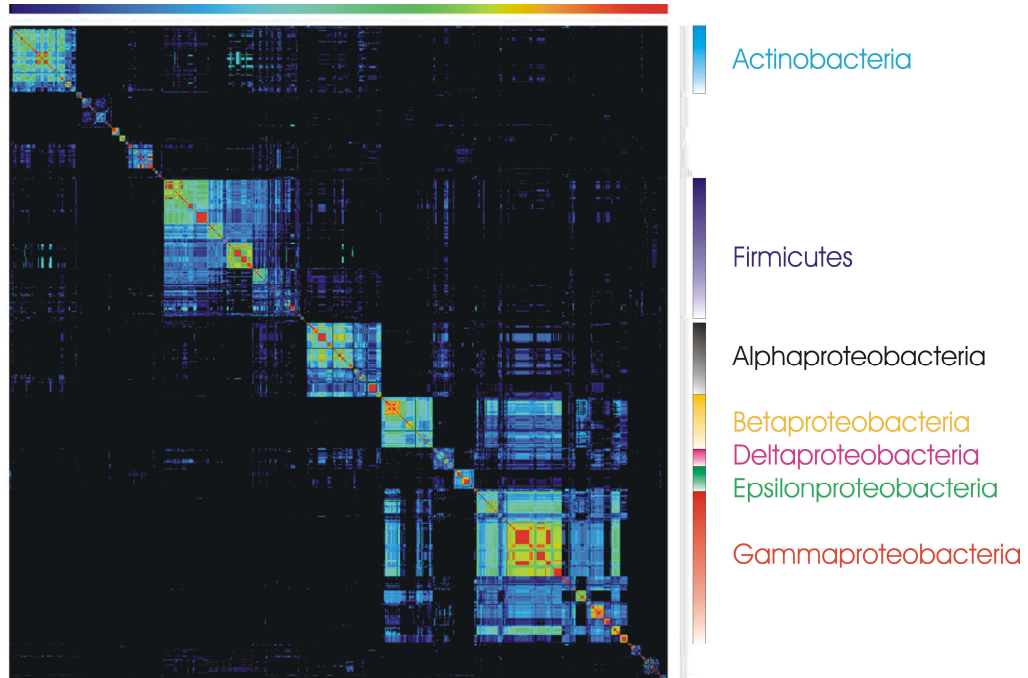


**Fig. 1.2** Comparison between the similarity values calculated for different classes. FL-FL: full length 16S rRNA as query and subject; PCR-FL: sub-sequences obtained by a virtual PCR on 16S rRNA genes as query and full length genes as subject; PCR-PCR: sub-sequence used as query and subject.

To isolate this specific signature, fasta.36 which allow the IUPAC encoding for degenerate nucleotides, was used considering primers sequences as query and 16S rRNA genes as reference database. A specific perl script was applied to the fasta36 output, checking annealing primers (which resulted for the 100% of sequences) and returning a multiFASTA of partial 16S rRNA “virtual amplicons”.

For data analysis, the result of BLAST research, was parsed using a perl script and organized in a square matrix sorted due to the Taxonomy browser of NCBI order (ncbi.nlm.nih.gov/Taxonomy). The same species were placed with the same entry order on each axis. Globally, this displacement could be considered an unrooted tree in which organisms are grouped without any temporal consideration or molecular clock evaluation.

The matrix was analyzed using MeV, a tool originally developed for microarray data analysis (Fig. 2). A similarity value between two sequences is represented by a dot in the matrix.



**Fig. 2.** Matrix visualization of the BLAST result. Each query is placed with the same entry order on each axis. The most abundant taxa are reported. Phyla: Firmicutes and Actinobacteria; classes Alpha-, Beta-, Delta-, Epsilon and Gammaproteobacteria.

Data were visualized considering a 82% of similarity as minimal average threshold. This value was determined considering the minimum normalized value observed between organisms belonging to the same phylum. The color gradient, from blue to red, indicates the increasing similarity.

Considering the threshold, all the main phyla can be distinguished, with an exception represented by Proteobacteria. Inside every phylum an increasing similarity can be observed in agreement with the closer relationship between species

- Actinobacteria phylum: the orange-to-red color of the main square, represents *Micobacterium* species.
- Firmicutes phylum: the *Bacillus* genus displays higher similarity values considering *Bacillus cereus* strains and *Bacillus thuringiensis* at the top left corner of the group. The middle group is represented by *Streptococcus* species: as expected, every sequenced strain of *Streptococcus agalactiae*, *Streptococcus pneumoniae*, *Streptococcus pyogenes* and *Streptococcus suis* can be distinguished. An higher diversity concerns *Bacillus* species, reflecting an higher heterogeneity probably related to the different lifestyle and type of sequenced organisms.
- Proteobacteria phylum: another scenario concerns members of Proteobacteria phylum. Only species belonging to the same class, a lower taxonomic level, originate well defined groups, while the imposed threshold allows the identification of all the other phyla. Similarity values, group together Betaproteobacteria (without the Comamonadaceae family) and Gammaproteobacteria (without the Chromatiales order) whereas only few species of Alphaproteobacteria seem to be related to the Beta-Gammaproteobacteria group and while Epsilonproteobacteria appears as an isolated clade.

Within Alphaproteobacteria, the Burkholderiaceae family displays common similarities, starting from 96%. Inside Burkholderia genus, different species of *Burkholderia cenocepacia*, *Burkholderia glumae*, *Burkholderia ambifaria* and *Burkholderia mallei* are characterized by s

similarity value higher than the 97.5%. Comamonadaceae form a well defined group which seems to be less related to the other Alphaproteobacteria.

- Considering the Enterobacteriaceae class, organisms belonging to the Escherichia genus can be distinguished starting from 98% of similarity. Moreover, a high value established between Escherichia and Shigella genera, underlining a historical taxonomic disquisition. Observing the Yersinia genus, a full similarity has been reported between *Yersinia pestis* and *Yersinia pseudotuberculosis*, a lower value for *Yersinia enterocolitica*.
- Vibrionaceae and Pseudomonadaceae seem to be closely related and some relationship can be observed between Vibrionaceae and Enterobacteriaceae families.
- Other taxa appear to be poorly related to each others. In this way, the orders of Xantomonadales, Legionellales and Cromatiales, that display low similarity values within the class but high similarity value and close relations if considered within the taxon itself.
- Previous reports state that, the entire Epsilonproteobacteria class does not display sufficient values that allow any putative linkage with other bacterial groups.

Interestingly, similarity values higher than the established threshold, can be observed in several inter-phylum comparison. Actinobacteria show a relation with the Firmicutes phylum if considering Bacillaceae and Streptococcaceae families. Moreover, another linkage was reported observing the same Firmicutes families and the Delta and Enterobacteriaceae class and family belonging to the Proteobacteria phylum.

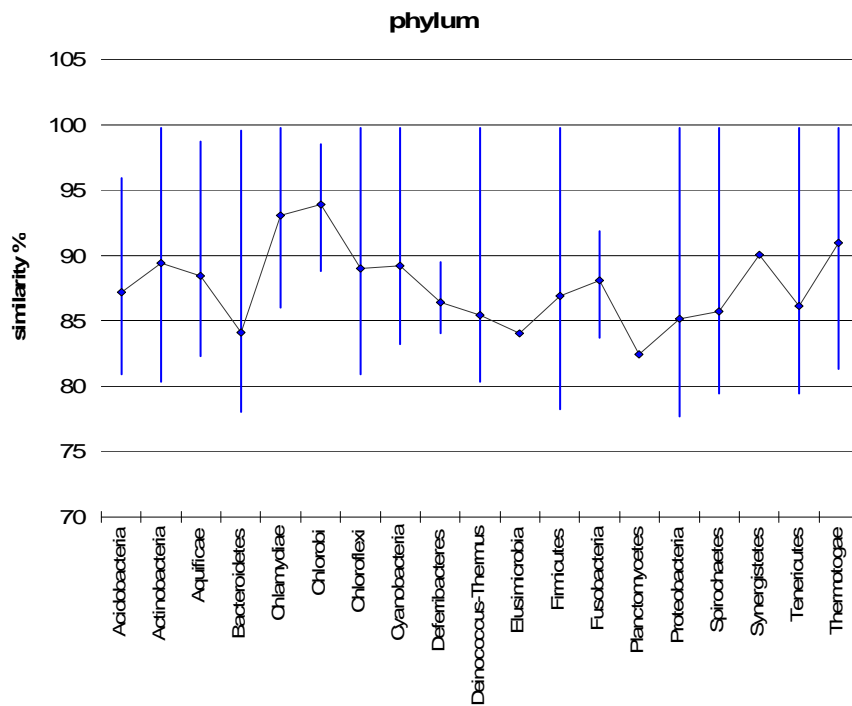
Result visualization in the color matrix shown, provided a wealth of relevant and novel information. Several cross-relations have been traced within each bacterial group, and heterogwony related to a comparison between a sub-sequence of 16S RNA genes and the entire gene was underlined.

The poor relationships have been established between Proteobacteria, for which only at class level an identification was possible: the inter-phyla result highlight

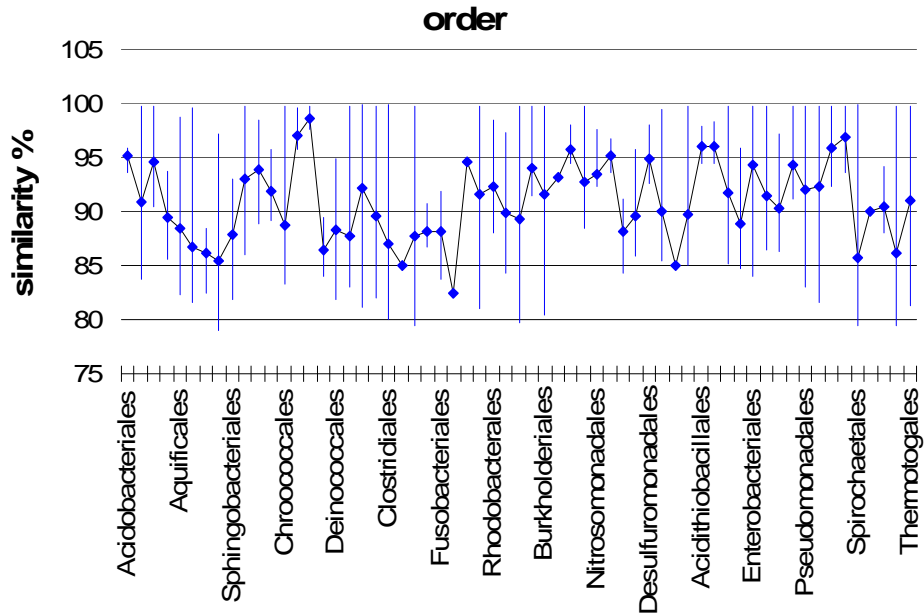
the need of further analysis.

To assess a numerical threshold for classification for each taxonomic level, obtained scores were graphically plotted considering maximum value, minimum value and average value for each taxon.

As observed in the matrix, the differences within the same phylum were higher. If a single value could be assessed for every taxonomic division, the intraspecific heterogeneity appeared remarkable. Average values were comprised between 85% and 90% (higher values concern homogeneous groups if sequenced type organisms are considered) and minimum and maximum values were indicative of a species (or few species) characterized by a divergent pattern with respect to the other.



**Fig. 3.** Similarities evaluated within the phyla considered. Blue dot represent the average similarity while the lines indicate the maximum and minimum calculated value.



**Fig. 4.** Average similarity evaluated at order level. 16 taxa are shown but the entire dataset counts 61 distinct orders.

The average values were higher with respect to the phylum and for some groups similarity was estimated around 90% (Fig. 3). At order level a lower similarity was observed between 90% and 95% and some groups shown a range of values which were less heterogeneous with respect to the others (Fig. 4).

Considering the average similarity between organisms belonging to the same order and to the same phylum, it is not possible to establish a number which can be used as universal threshold.

A better resolution and a numerical definition, could be assessed considering the similarity values in a genus level comparison (Fig. 5).

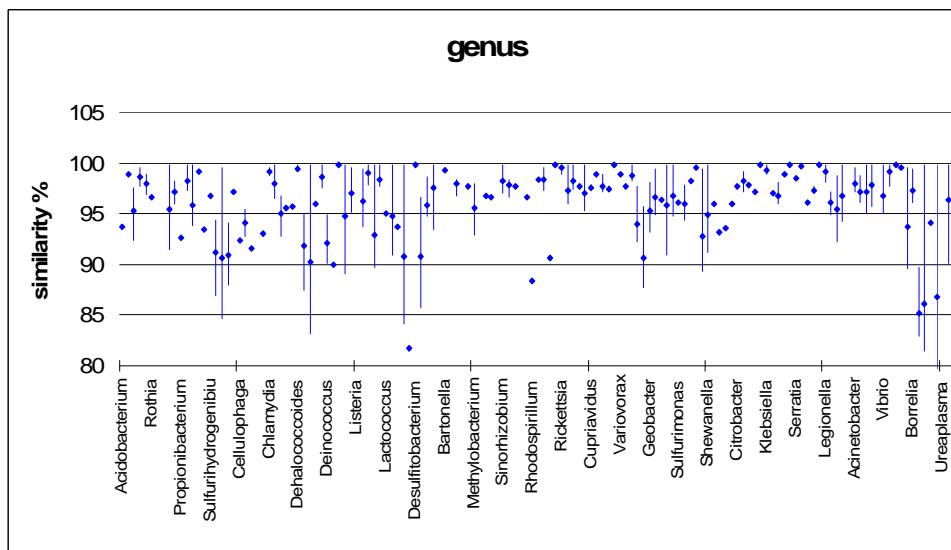
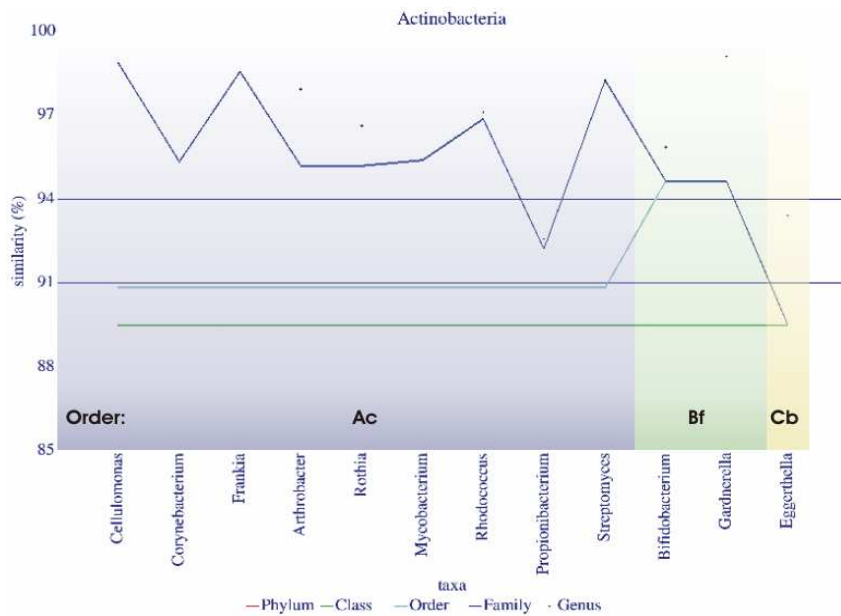


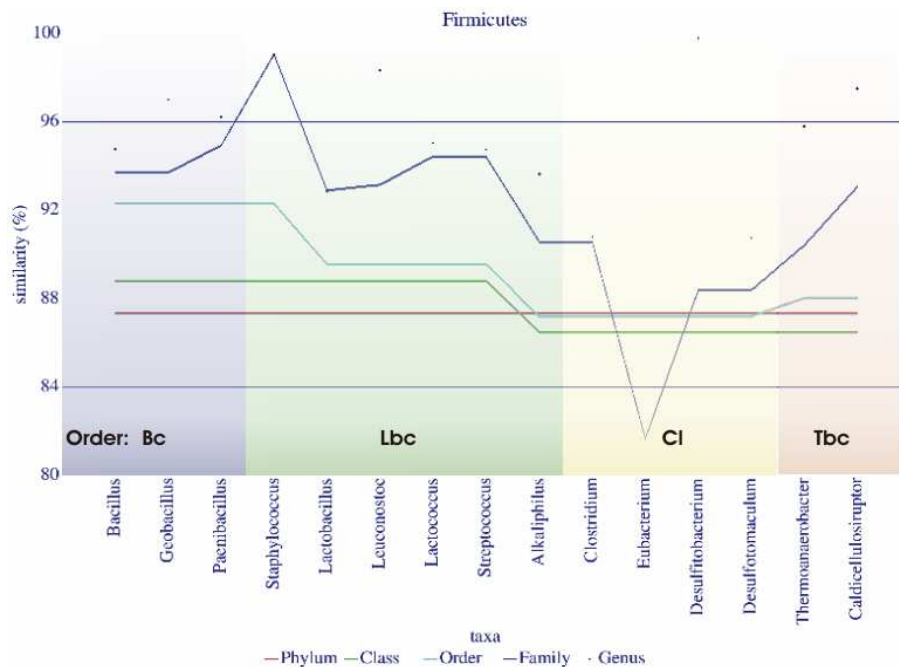
Fig. 5. Average similarity evaluated at genus level. In this graph, 143 genus are plotted.

For a wide amount of organisms, average value was located between 95% and 99% of similarity. This results, indicates that organisms belonging to the same genus, were rather similar, allowing several considerations about this threshold in determining and supporting a sequence classification. Other informations have been obtained considering the unique average value evaluated for each phylum, class, order, family and genus.

If an average value was reported for a phylum, similarity within the single class, orders, family and genus was compared together. The same similarity value was evaluated for phylum and class levels within the Actinobacteria-related organisms (Fig. 6).

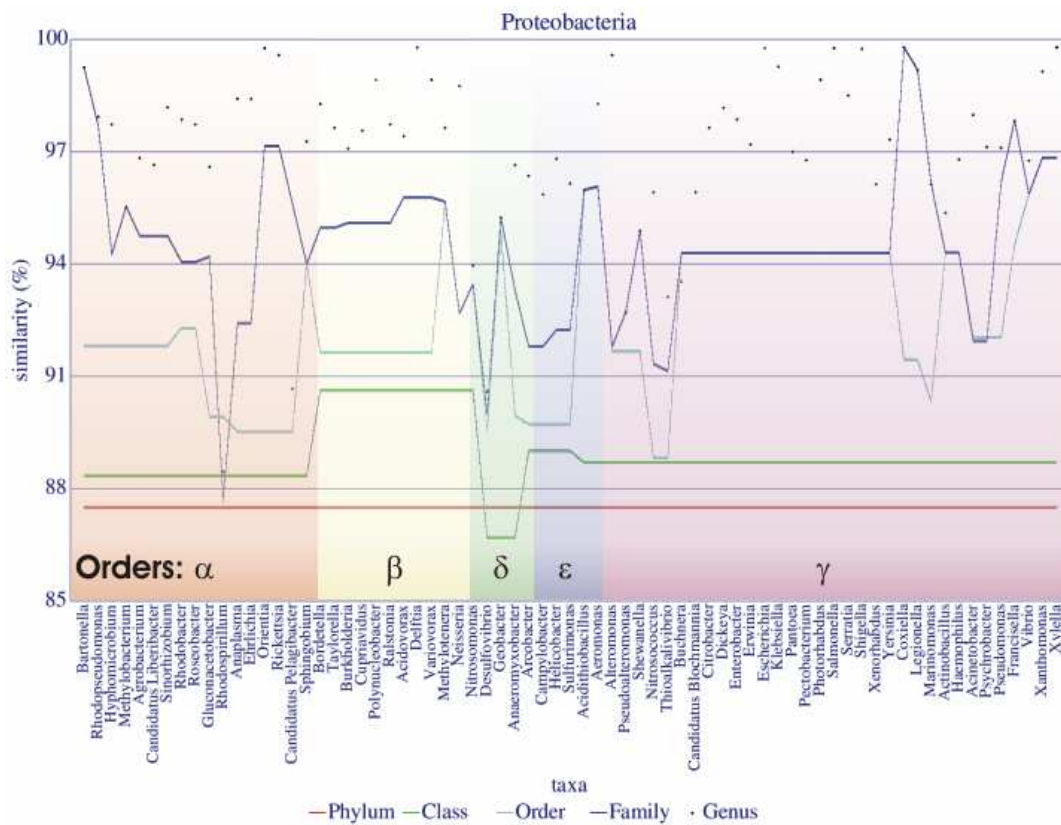


**Fig. 6.** Average similarity values within the Actinobacteria phylum. Three orders are shown: Actinobacteriales (Ac), Bifidobacteriales (Bf) and Coriobacteriales Cb.



**Fig. 7.** Average similarity values within the Firmicutes phylum. The orders shown are: Bacillales (Bc); Lactobacillales (Lbc); Clostridiales (Cl); Thermoanaerobacteriales (Tbc). Eubacterium seems to be poorly similar with respect to the other Firmicutes genera.

Within the Firmicutes phylum a decreasing similarity can be observed considering lower taxonomic levels of the Straphylococcus- and Lactobacillus- related organisms. Lower values have been reported within the Clostridiales order: considering average value, the inter-similarity resulted was lower with respect to the value reported for the entire phylum (Fig. 7).



**Fig. 8.** Average similarity values within the Proteobacteria phylum.

Within Proteobacteria an heterogeneous situation was observed (Fig. 8). In particular, an increasing similarity between sequences was defined for all the Proteobacteria from phylum, class (exception for Deltaproteobacteria) and orders. At genus level all these organisms were grouped between 96% and 100% values.

## 5.2 Cluster analysis and sequences annotation

Starting from considerations emerged during the data analysis about the sequencing result of ten bioreactors defined in chapter 4, par. 4.4.1, several programs were written for the data annotation.

A first annotation was carried out considering a sequence collection downloaded from the RDP site and represented only by high quality and fully annotated 16S rRNA genes, a total amount of 7425 sequences. Considering a sample where a specific Planctomycetes-related pathway was observed, no sequences related to this taxon were found.

To expand the reference dataset, CAMERA databases were considered. An enormous amount of data derived from several metagenomics and genomics sequencing projects are updated in this repository.

Database characteristics and size are reported in Table 1.

Name	Description	no. bases	no. sequences
NCBI RefSeq Microbial Genomes	All NCBI RefSeq microbial genomes sequences. This data set is created from genomic sequences under the microbial section in NCBI RefSeq	13.101.019.356	341.346
All Prokaryotic Genomes	Genomic sequences from publicly available draft and finished prokaryotic genomes. This data set included prokaryotic genomes sequences in both GenBank and RefSeq	29.948.347.987	3.795.052
Non identical nucleotide sequences	Non-redundant nucleotide database	255.115.078.617	78.882.301

**Tab. 1.** Summary information on reference databases used

The nucleotide BLAST results, here defined as primary annotation, shown an higher variability with respect to the RDP-based analysis, returning a large amount of unidentified sequences (and annotated as “no\_taxon”).

A cluster analysis considering the 97.5% of similarity as threshold was carried out using cdHIT. This value is the generally established value for species definition.

A differing result between clustering and classification was detected using perl and MySQL as informatics tools. For classification, Tax\_ID which univocally recognized each taxonomic node, was considered.

**Primary annotation (1)**

Query_ID, 97%	Phylum	Class	Orders	Family	Genus	Species
G2HZY5308I7H0I	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	no_taxon	uncultured ammonia-oxidizing bacterium
G2HZY5308I7DXB	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	Nitrospira	uncultured Nitrospira
G2HZY5308JB3XZ	no_taxon	no_taxon	no_taxon	no_taxon	no_taxon	uncultured bacterium
G2HZY5308JHSHP	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	Nitrospira	Nitrospira sp.
G2HZY5308JGUL7	Proteobacteria	Betaproteobacteria	no_taxon	no_taxon	no_taxon	uncultured beta proteobacterium
G2HZY5308JEFZC	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	Nitrosomonas	uncultured Nitrosomonas sp.
G2HZY5308JL4G5	no_taxon	no_taxon	no_taxon	no_taxon	no_taxon	uncultured bacterium

**Principle of coherence (2)**

G2HZY5308I7H0I	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	no_taxon	uncultured bacterium
G2HZY5308I7DXB	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	no_taxon	uncultured bacterium
G2HZY5308JB3XZ	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	no_taxon	uncultured bacterium
G2HZY5308JHSHP	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	no_taxon	uncultured bacterium
G2HZY5308JGUL7	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	no_taxon	uncultured bacterium
G2HZY5308JEFZC	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	no_taxon	uncultured bacterium
G2HZY5308JL4G5	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	no_taxon	uncultured bacterium

**Final annotation (3)**

G2HZY5308I7H0I	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	no_taxon	uncultured ammonia-oxidizing bacterium
G2HZY5308I7DXB	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	Nitrospira	uncultured Nitrospira
G2HZY5308JB3XZ	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	no_taxon	uncultured bacterium
G2HZY5308JHSHP	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	Nitrospira	Nitrospira sp.
G2HZY5308JGUL7	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	no_taxon	uncultured beta proteobacterium
G2HZY5308JEFZC	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	Nitrosomonas	uncultured Nitrosomonas sp.
G2HZY5308JL4G5	Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	no_taxon	uncultured bacterium

**Fig. 9.** Several annotations were found for these sequence clusters at 97.5%, at each taxonomic level (1). Following the ‘principle of coherence’, annotation were maintained whereas identical among queries and extended to the unidentified sequences (2). As a final step, primary and secondary annotation were considered as one (3).

In Fig. 9, different taxonomic attributions were found for groups of sequences clustered at 97.5% of similarity as threshold.

A subsequent analysis on the output files showed that because of the massive amount of data contained in each database, few differences between sequences can determine a different hits matching.

To solve this problem, a ‘principle of coherence’ was applied in a two-steps procedure (in collaboration with Dr. Nicola Vitulo).

- Annotation results (‘primary annotation’) were maintained for a taxonomic level whereas conserved among sequences and extended just to the relative ‘no\_taxon’ annotated sequences. A ‘no\_taxon’ classification or ‘uncultured bacterium’ for the species, were applied for all the sequences when

different taxa were recognized by primary annotation (e.g. genus and species, step '2', Fig. 9). This annotation strategy, is strongly supported but the loss of information is remarkable for lower taxonomic levels.

- Primary annotation and clustering approach were considered together. Following the same 'principle of coherence', sequences were annotated while the primary annotation remains unchanged extending it to the 'no\_taxon' queries. When discrepancies were found, only primary annotation was allowed (e.g. genus and species, step '3', Fig. 9).

The resulting annotation on more than 60.000 sequences analyzed was graphically summarized in Figs. 10, 11, 12.

Number of 'no\_taxon' annotated sequences has been reported for each step of annotation:

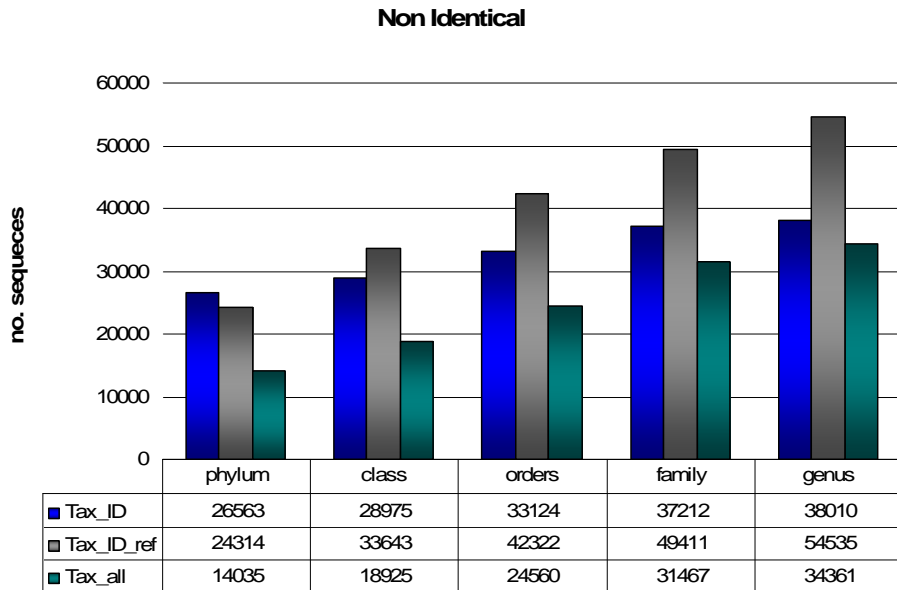
- Tax\_ID identified the primary annotation.
- Tax\_ID\_ref: for the first cluster analysis based only on the principle of coherence.
- Tax\_all: which interpolates the two data.

The 30% of not annotated sequences was found in each sample if primary annotation was considered. This value increased to more than 45% for higher taxonomy level.

The possible presence of unknown novel organisms was indicated: not annotated queries were grouped in specific clusters for the considered samples. This observation was supported considering the specific experimental conditions imposed for the development of a poor but defined consortium of microorganisms.

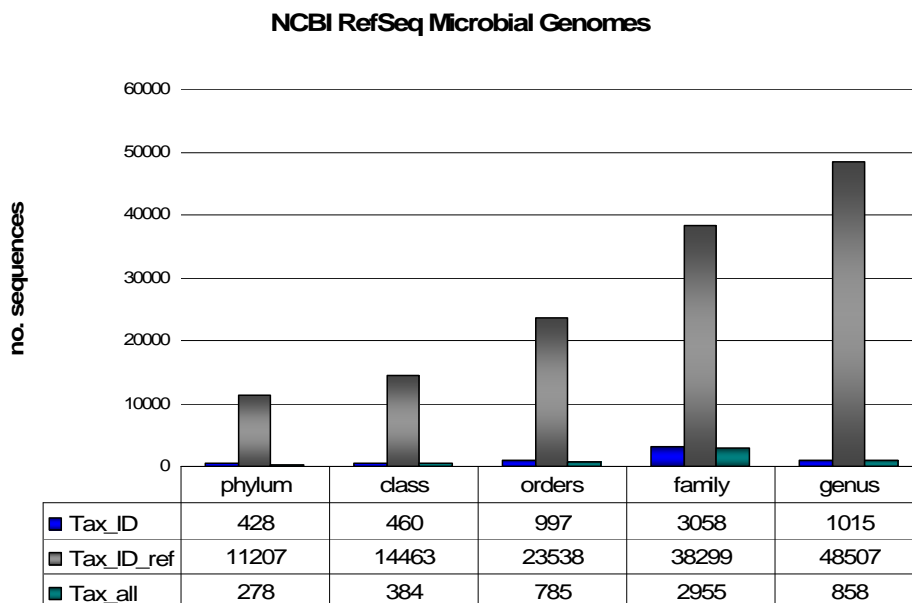
A correlation could be observed by the evaluation of the three reference databases.

The higher number of not annotated sequences was found using the Non\_identical as reference database, derived by the non-redundant database of NCBI (Fig. 1.1).



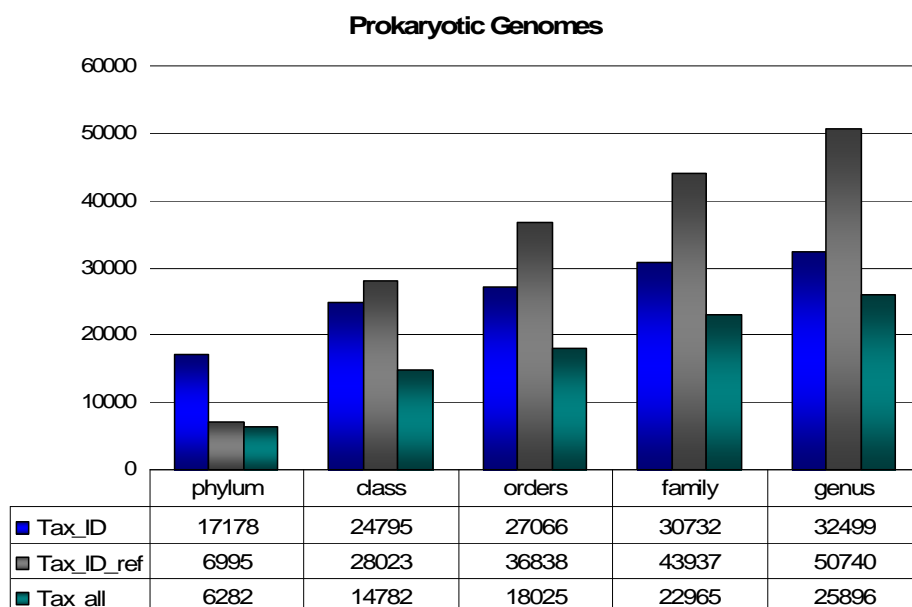
**Fig. 10.** Number of sequences identified as ‘no taxon’: Non identical dataset.

NCBI\_RefSeq microbial genomes derived from the microbial section of NCBI. A better annotation was obtained but comparing it with cluster analysis an higher diversity could be observed at family and genus level. Infact, after the first step for which a Tax\_ID\_ref was assessed, an increasing number of sequences reported as “no\_taxon” has occurred, underlining the poor coherence between the identified labels (Fig. 1.2).



**Fig. 11.** Number of sequences identified as ‘no taxon’: NCBI RefSeq Microbial Genomes.

Finally, data derived from completely and draft sequenced bacterial genomes that characterize Prokaryotic\_genomes databases was compared. It could be considered the most interesting because uploaded data derives from single projects for which other biological, physiological, biochemical, structural informations may be obtained.



**Fig. 12.** Number of sequences identified as 'no taxon': Prokaryotic genomes.

Observing the different results, for each database and for each taxonomic level a decreasing number of not annotated sequences has been reported, increasing the information relative to the biodiversity of the specific sample.

### 5.3 Conclusions

Sequences comparison with annotated subjects represents the best way for bacterial identification and one fundamental step for a classification in which all the species are collected following a logic and historical path.

The huge amount of differences within and between microbial taxa were underlined walking instead the 16S rRNA genes analysis and starting from classified bacteria to define similarity values.

Taxa comparison revealed that the huge biodiversity that characterizes the Proteobacteria phylum was reflected by ribosomal genes comparison, underlining a lower value with respect to the other microbial phyla. The upgrade to 'phylum' rank suggested for the Proteobacteria classes, the low heterogeneity that characterizes the Actinobacteria phylum and other examples, underline the complexity and the difficulty in finding a standard protocol for bacterial identification and classification.

## **Chapter 6. Anammox bioreactor: results and discussion**



Ultrastructural observations of the anammox-like microbiota and the identification of putative Planctomycetes, were coupled with a PCR-based approach.

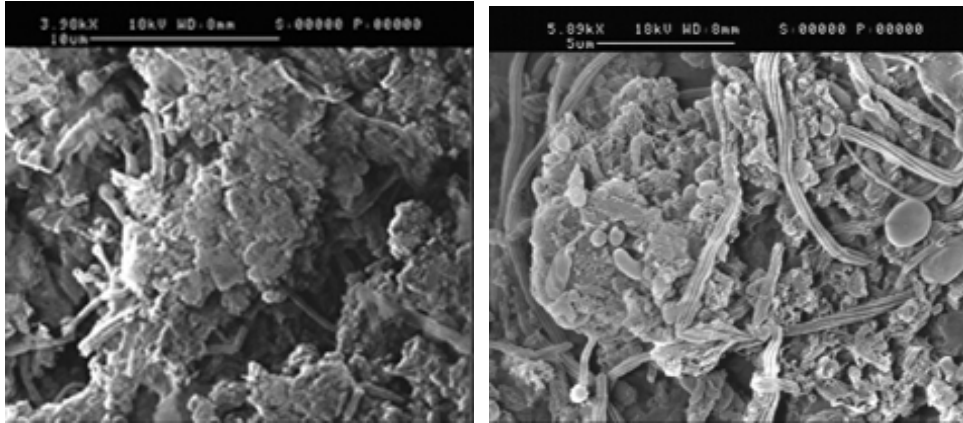
### **6.1 Electron microscopy**

Electron microscopy was applied to analyze the Anammox sludge and its ultrastructural features from samples taken from the Eurotec WTT bioreactor.

Scanning electron microscopy revealed a granular sludge in which several unidentified filamentous structures appeared interspersed in a granular matrix.

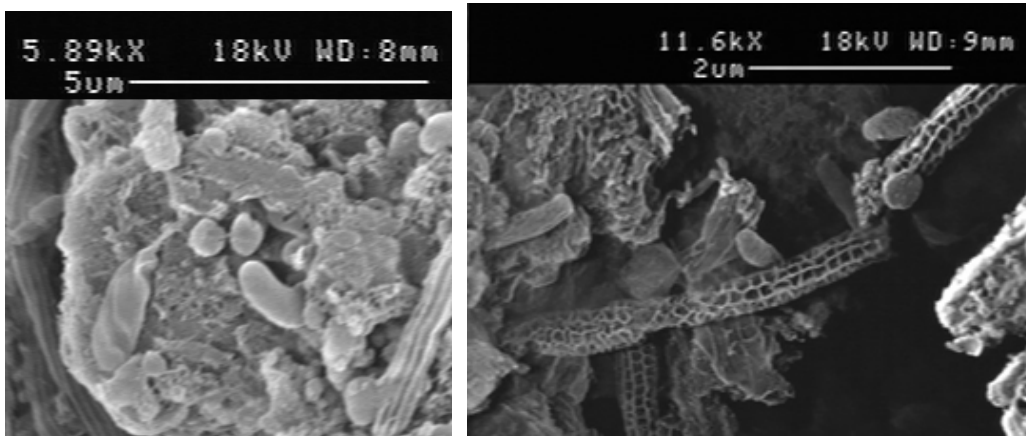
Microbial coccoid cells and other morphological types were detected revealing a polymicrobial community.

### 6.1.1 Scanning Electron Microscopy observation



**Fig. 1.** SEM images of the amx sample. Bacterial cells populate a granular matrix characterized by unidentified filamentous structures.

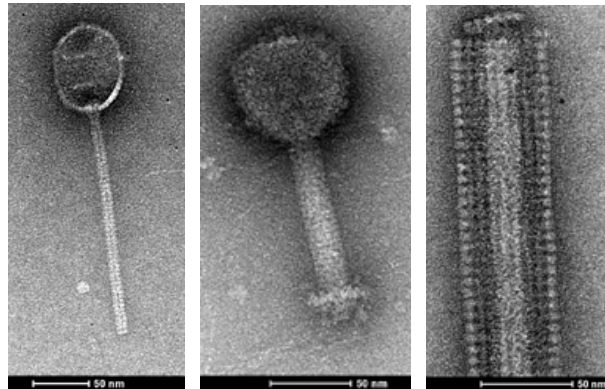
The images displayed was obtained from a Eurotec WTT sample. The bioreactor seems to be characterized by a granular matrix (Fig. 1) in which several cellular types are visible, coccoids or with extended shapes (Fig. 2). A huge amount of filamentous structures was detected but their nature was not deciphered.



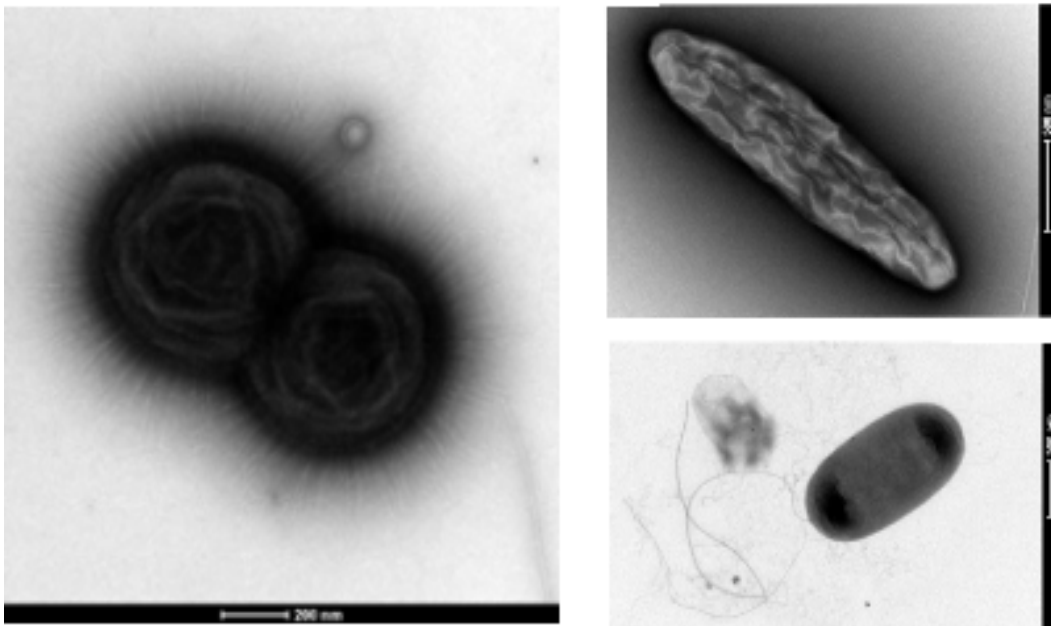
**Fig. 2.** Several cellular types were observed. A huge amount of linear, alveolar fibers characterized the amx sample. Their identification remains uncertain.

### 6.1.2 Negative Staining Electron Microscopy

Thanks to the negative staining microscopy, a vast amount of different microorganisms that populate the Anammox bioreactor was discovered. The presence of several cellular types and phages, could suggest a complex and complete community structure.



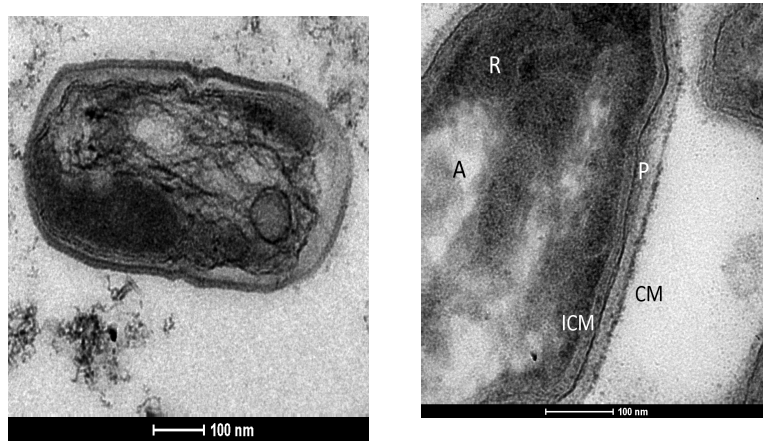
**Fig. 3.** Phages and other viruses detected in the Anammox samples. Scale: 50 nm



**Fig. 4.** Negative staining electron microscopy. Other specific features of the cellular types previously observed. Cocci and different types of bacilli, motile and not motile. Scales: 200 nm (left), 500 nm (right).

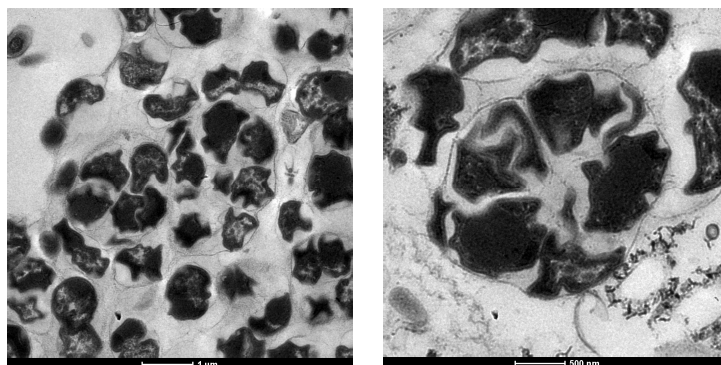
### 6.1.3 Transmission Electron Microscopy

Confirming the existence of a putative Anammox-type consortium, ultrastructural informations related to Planctomycetes bacteria were provided by transmission electron microscopy. All cellular features that characterize this particular phylum and Anammox bacteria were apparently identified: an inner compartment that could be the anammoxosome, riboplasm, paryphoplasm separated from the riboplasm by an intracytoplasmic membrane and the cytoplasmic membrane.



**Fig. 5.** TEM images from anammox bioreactor. A compartmentalization is clearly visible. An hypothetical Planctomycetales cell structure is shown: anammoxosome (A), riboplasm (R), paryphoplasm (P), intracytoplasmic membrane (ICM), cytoplasmic membrane (CM). Scale: 100 nm

Anammox granules were also identified. Some cell shape alterations (shrinking) could have occurred in the preparation of the sample for TEM observations, during the dehydration step. Moreover, an extracellular matrix could be observed possibly representing a biofilm enclosing each cell group.

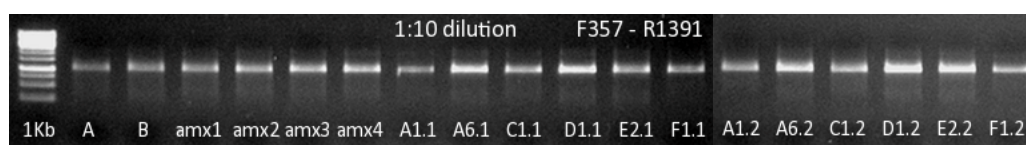


**Fig. 6.** TEM images from anammox bioreactor. Hypothetical anammox granules are shown. Scale: 1 μm (left), 500 nm (right).

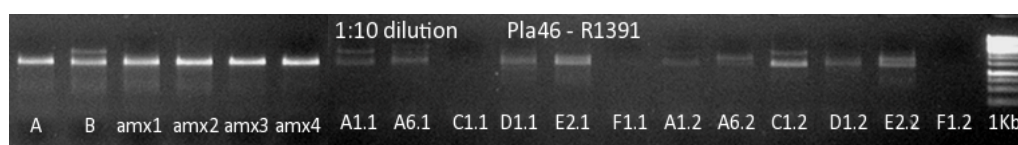
## 6.2 Semiquantitative PCR reaction

Following development and evolution of the microbial community, PCR with universal (F357-R1391) and Planctomycetes-specific primers (Pla46-R1391) was carried out.

All samples positively amplified when using universal primers (Fig. 7). Moreover a signal which may suggest an Anammox-enriched population was clearly detected for all the EurotecWTT samples (Figure 8).



**Fig. 7.** F357-R1391 universal primers PCR.

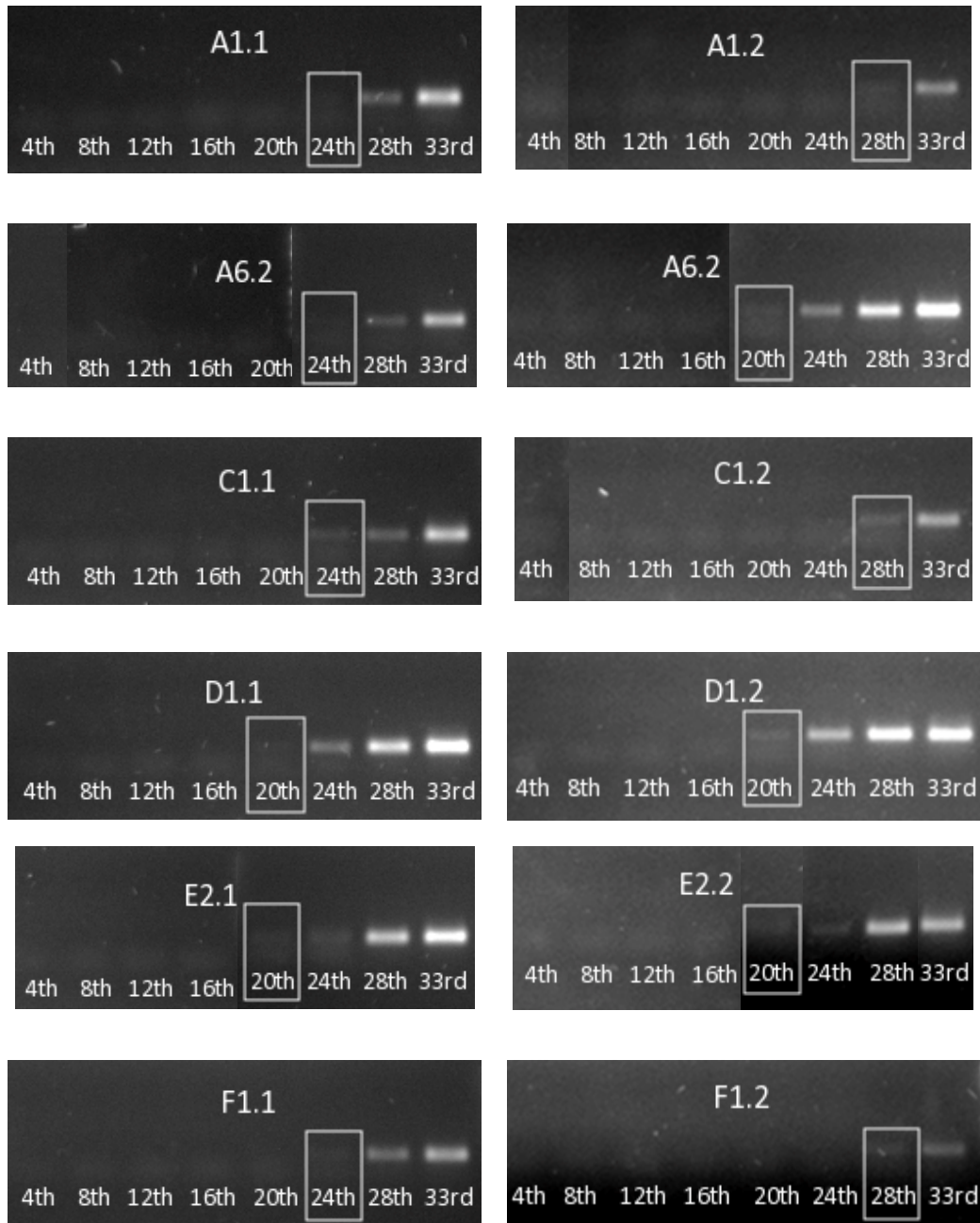


**Fig. 8.** Pla46-Planctomycetes specific primer and R1391-universal primer PCR. Eurotec WTT samples amx1-amx4 and several Milan bioreactors appear to have a Planctomycetales population,

Semiquantitative PCR may be a useful technique to obtain a preliminary information about the microbial population analyzed, Notwithstanding the possible occurrence of multiple 16S rRNA operons in bacterial genomes and the presence of inhibitory substances which can affect a specific sample .

A semiquantitative approach was followed to track the microbial population development dynamics. in the different samples.

Semiquantitative PCR was carried out by universal primers F527-R790 (Fig. 9) and normalized DNA amounts (showed, 0.66 ng/ul) to evaluate and compare differences in total bacterial content. Bioreactors at different time points were analyzed.



**Fig. 9.** F527-R790 semiquantitative PCR on a normalized template dilution (0.66 ng/□). A more abundant microbial population is indicated by a band after 20 cycles for A6.1, D1.2, E2.2 samples. The appearance of a signal after 24 cycles is indicative of a smaller community for A1, C1 and F1 samples.

A visible band after fewer PCR cycles was detected for samples with more abundant microbial populations: both the time points of A6, D1, E2 samples displayed a band after 20 cycles, earlier with respect to the others, while a different scenario was recorded for samples A1, C1 and F1. The appearance of a signal between 24 and 28 PCR cycles, is indicative of a smaller bacterial amount.

For most of the templates, an earlier visible signal was shown comparing the first time points with the second time points (A1, D1, F1 samples): a more limited number of bacterial species was supposedly selected by cultural conditions.

Sample A6 and D1 displayed an opposite trend while in sample E it was not possible to define the threshold due to variations in all the experimental replicates. Optimal growth conditions could be assessed for pre-existing bacteria in A6 and D1 samples, determining an increasing growth rate in such populations.

For sample A and sample B, a semiquantitative PCR was carried out using Anammox-specific primers F818-R1064. Results are reported in Fig. 10.a.



Fig. 10. a.

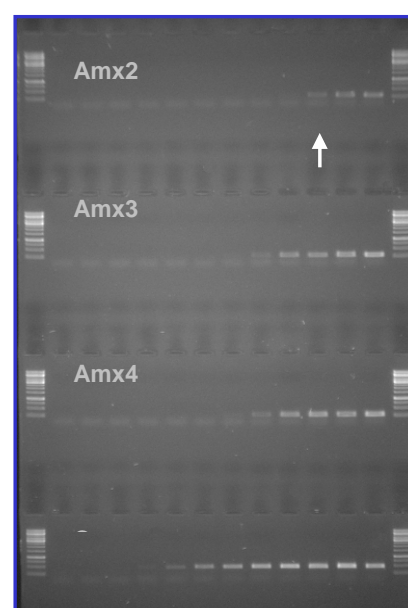


Fig. 10. b.

**Fig. 10.a., 10. b.** F818-Planctomycetes specific primers and R1064-universal primers on A and B Milan samples (Fig. 10.a). A band after 27 rather than 30 cycles indicates an higher amount of Planctomycetes-related bacteria for sample A than sample B. Fig 10.b. Planctomycetes specific primers on amx samples.

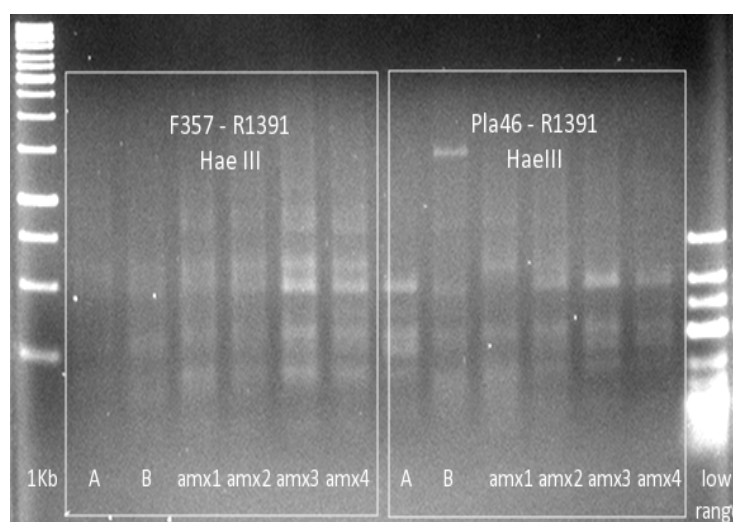
An higher amount of Anammox bacteria was clearly observed in sample A when compared with sample B: a poorer amount of these bacteria in the Milan bioreactor was suggested by a signal around 27 and 30 PCR cycles respectively, indicating that Anammox represent only a small fraction of the entire microbial community in those samples.

The same experiment was carried out on material from the Eurotec WTT

bioreactor in which an increasing Anammox-like activity was recorder during time for amx2, amx3 and amx4 samples (Fig. 10.b.)

### 6.2.1 Restriction Fragment Length Polymorphisms analysis

Microbial communities were analyzed using a RFLP-based approach. The definition of a specific pattern was allowed using restriction enzymes and several differences were observed considering biodiversity and environmental conditions. Considering amx amplicons from EurotecWTT and A, B samples, a general pattern was obtained using bacterial F357-R1391 universal primers and *HaeIII* as restriction enzymes (Fig. 11, left side) .

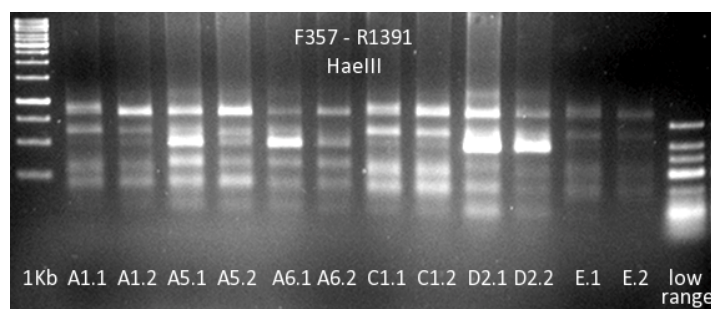


**Fig. 11.** RFLP on F357-R1391 universal primers and Pla46-Planctomycetes R1391-universal primers amplicons digested with *HaeIII* endonuclease.

A more heterogeneous pattern was defined for sample amx1 which represents the Anammox community starting inoculum. An homogeneous condition was detected for other samples indicating a decreasing and stabilized community along with the progress of incubation: a similar pattern, clearer and well defined for subsequent time points was observed for the amx3 and amx4 sampling.

A corresponding trend was recognized due to Planctomycetes amplification using phylum specific primer Pla46 and universal primer R1391. (Fig. 11, right side)

The development of the microbial populations, was assessed also for six pilot batches set up in Milan by care of Prof. F. Adani research group (Fig. 12.).



**Fig. 12.** RFLP on different F357-R1391 universal primers amplicon digested with *HaeIII* endonuclease.

Interestingly, a conserved pattern can be evaluated for each time point of the same sample and for similar growth conditions. This is true for samples A5 and A6, which derive from A1 but are characterized by glycerine supplementation to evaluate the effect of organic carbon on microbiota. Different conditions were evaluated for two time points of the C1 experiments in which the effect of doubled carbon to nitrogen (COD/N=2) was monitored.

Finally, a distinct pattern was observed also for sample D2 in which a population of denitrifying bacteria was surely present as it was deliberately inoculated.

The general assumption for which culturing conditions can determine different selective pressures stimulating specific organisms and consortia, was assessed by these experiments with the identification of a close relationship between growth conditions and resulting RFLP.

## 6.3 Sequencing

### 6.3.1 Sanger sequencing

In first instance Anammox bacteria were detected using a Sanger sequencing approach.

Total DNA was extracted by EurotecWTT amx3 sample and amplified using the F818 anammox specific primer and R1064.

Without cloning, a clean electropherogram was obtained: nucleotide BLAST

research indicated *Brocadia anammoxidans* as the best candidatus due to a 97% sequence similarity.

A survey was carried out using a nonredundant ad-hoc database of 98 Anammox bacteria 16S rRNA genes. Subsequences corresponding to the amplified region were obtained due to a fasta35 search and perl scripts, aligned using ClustalW2 and a data visualization was obtained using GBLOCK. Specific analysis of this rRNA gene fragment, revealed high conservation features while the estimated similarity value did not enable a clear species identification within Anammox bacteria.

### 6.3.2 Roche-454 sequencing

Ten samples were selected on the basis of biochemical results, semiquantitative PCR and RFLP for a 16S rRNA gene amplicon library tag-sequencing by Roche GS-FLX 454 Titanium Upgraded Genome Sequencer .

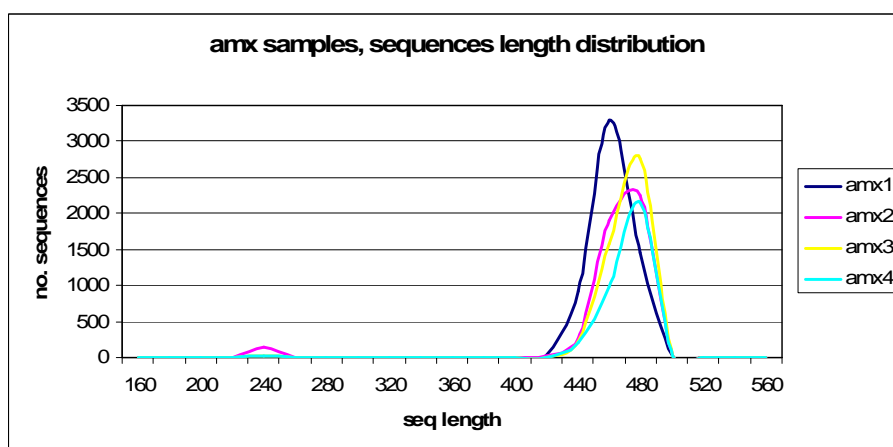
Summary information on the 61365 sequences obtained are reported in Tab. 1. Sequences were trimmed at the Galaxy portal on the basis of length (>200 b) and quality values at 5' and 3'. Moreover, also bases with a quality value lower than 20 were depleted.

Sample	Number of reads	Number of reads after quality and length trimming	Average of read length after quality and length trimming	Shortest reads	Longest reads
amx_1	5730	5724	448	208	498
amx_2	4606	4599	451	213	498
amx_3	4639	4633	458	208	520
amx_4	3398	3397	458	212	503
A1.1	6307	6298	458	297	658
A1.2	1791	1791	457	210	670
A5.1	14562	14539	461	209	506
A5.2	6818	6812	462	201	502
C1.1	7811	7800	457	228	499
C1.2	4774	5766	456	228	504

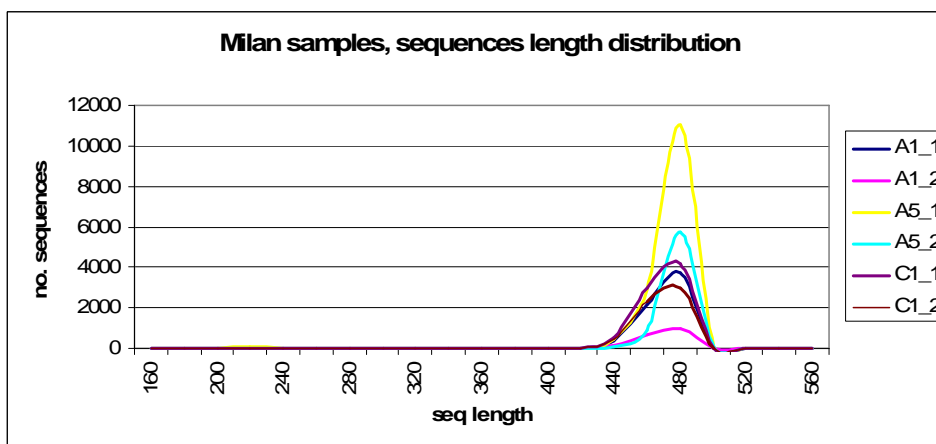
**Tab. 1.** Summary informations on 454 sequences obtained.

Numbers of sequences were rather homogeneous among samples. An exception was represented by sample A1.2 with 14562 reads and sample A5.1 with 1791 reads. It is possible that an error occurred during amplicons quantification determining an unbalanced amount of pooled PCR products.

The length of most sequences was comprised between 440 and 480 bases as expected, a distribution length graph was obtained for each sample to examine whether the association between primers and MID could affect the efficiency of sequencing (Fig.s 13-14).



**Fig. 13.** Sequences length distribution of amx samples. Most sequences have a length comprised between 440 and 480 bases.



**Fig. 14.** Sequences length distribution in the samples from the Milan bioreactors.

### 6.3.3 Cluster analysis

Cluster analysis using CdHIT was carried out on sequences on the basis of three similarity values, 100%, 99% and 97.5%. The number of cluster, is summarized in Tab. 2.

<b>SAMPLE</b>	<b>Number of clusters 100%</b>	<b>Number of clusters 99%</b>	<b>Number of clusters 97.5%</b>
Amx_1	4889	2189	1536
Amx_2	3409	1323	900
Amx_3	3126	1094	743
Amx_4	2355	515	278
A1.1	3573	540	265
A1.2	1155	278	163
A5.1	7732	891	318
A5.2	3633	551	237
C1.1	4236	756	385
C1.2	3307	696	398

**Tab. 2.** Summary information on the number of clusters obtained.

Considering 100% of sequence similarity, the number of unique reads can be obtained. The number of clusters for amx1 was higher with respect to the equivalent value established for the other samples. This sample, represents the first bioreactor inoculum originated from a nitrification-denitrification sludge in which a complex nitrification-denitrification-anammox consortium may proliferate. Biodiversity of such environment is more heterogeneous if compared with a bioreactor microbial population in which the growth conditions are constantly monitored.

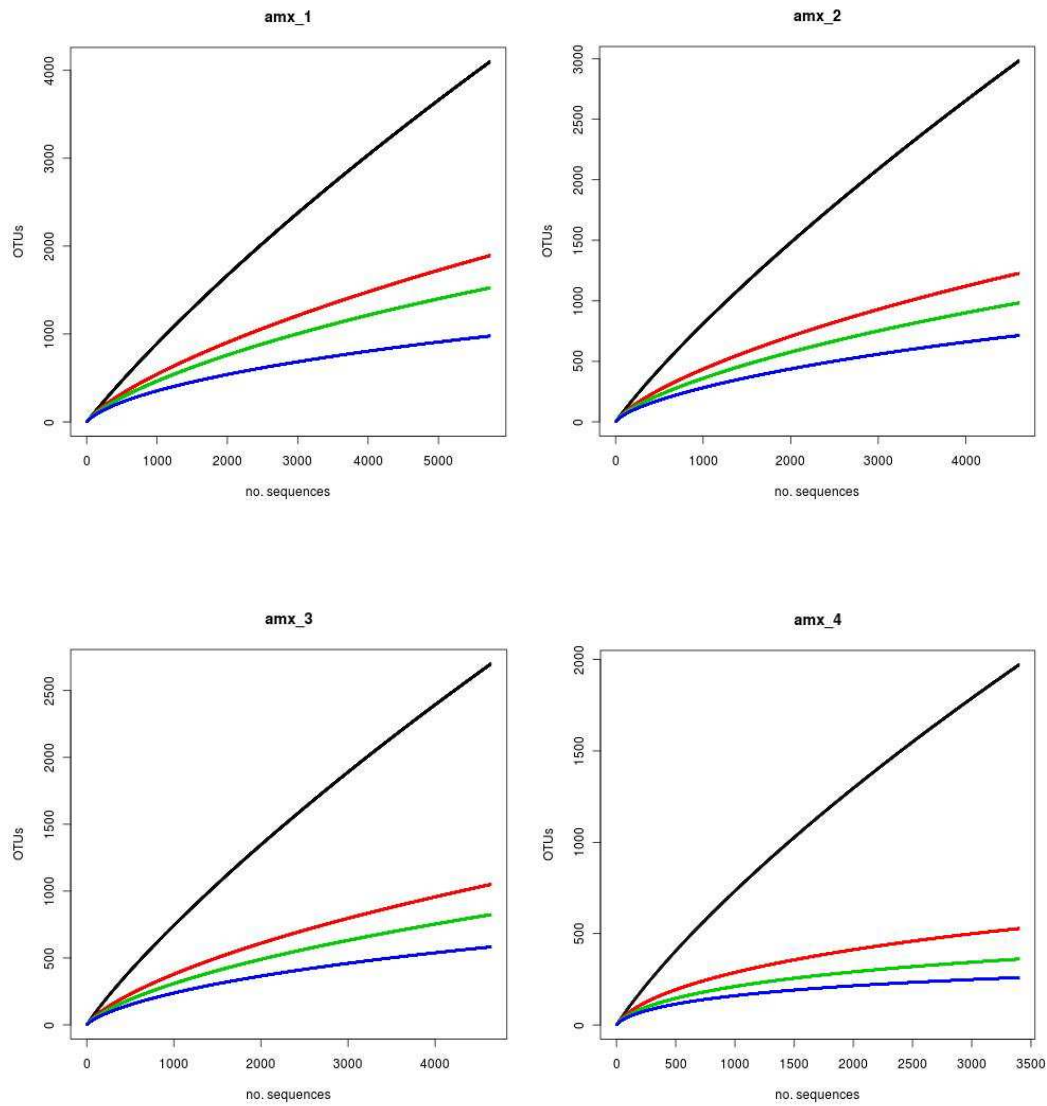
A 99% cutoff of similarity takes into consideration possible sequencing errors, intraspecific microvariability related to different 16S rRNA genes within the same genome or different microbial populations. The resulting number of clusters was less than half..

A 97.5% of similarity represent the common accepted threshold for species definition. The number of clusters, theoretically, indicates the number of species which populate an environment. Several considerations about bacterial species definition have been just reviewed, and interpolating these data with the annotation results seems to represent a good strategy to improve classification at every taxonomic level.

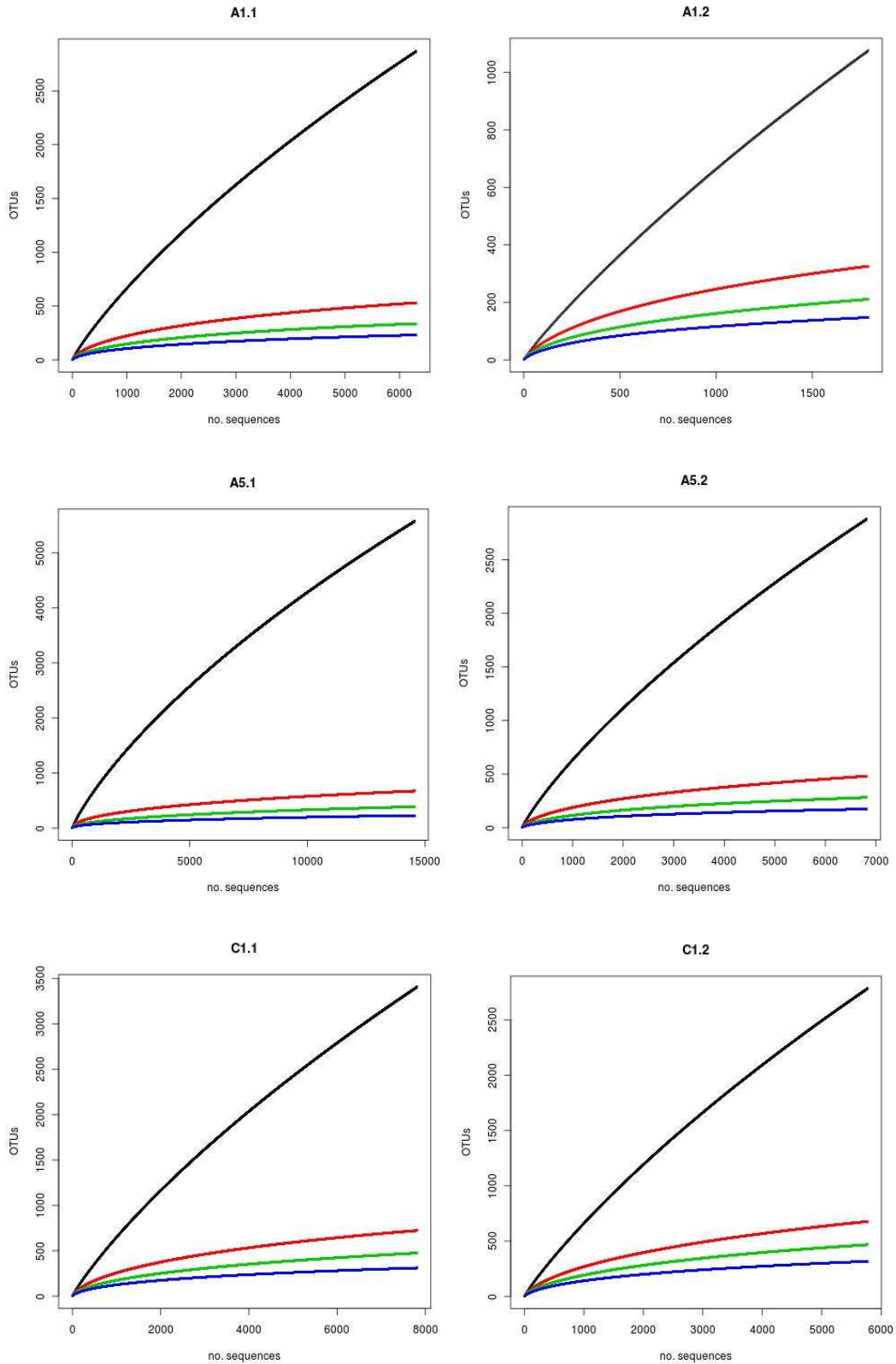
A decreasing number of clusters was observed, consistent with the timewise incubations starting from the first sample amx1 (richest in diversity) and proceeding through amx2 to amx3 and amx4. For the Milan batches values higher values for the A5 and C1 samples may indicate a possible regrowth of species underrepresented in the Anammox microbial population.

### 6.3.4 Rarefaction analysis

All reads were subjected to rarefaction analysis to investigate to which extent the sequences obtained could be sufficient to cover the entire microbial biodiversity (Fig. 15.a, 15.b).



**Fig. 15.a.** Rarefaction curve on amx samples. Different lines represents the different number of OTUs evaluated considering distinct clustering thresholds (black: unique; red: 0.03; green: 0.05; blue: 0.10)



**Figure 15.b.** Rarefaction curve on Milan bioreactors sequences. Different lines represents the different number of OTUs evaluated considering distinct clustering thresholds (black: unique; red: 0.03; green: 0.05; blue: 0.10)

A particular distribution was obtained using MOTUR. To trace a rarefaction analysis, sequences were clustered in OTUs following a decreasing similarity value: 100% or unique (black lines), 97.5% or 0.01 (red lines), 95% or 0.3 (green lines) and 90% or 0.1 (blue lines). Unique sequences, did not reach a plateau, however the variability within this pool could be in large part due to routine sequencing errors inherent to the process. In fact a dramatically dropping trend and a near-plateau shape of the curves was observed when settling for lower similarity values. In particular, as reported by, sequencing errors may affect a cluster-based analysis and a wide diversity in the number of clusters can be obtained if 100% or 99% are considered as similarity thresholds. Thus without considering ‘unique’ sequences, it appears that the number of reads achieved is sufficient to obtain an acceptable richness coverage for some samples (Table 3).

Sample	Reads	0.03		0,05		0,10	
		OTU	ACE (lci, hci)	OTU	ACE (lci, hci)	OTU	ACE (lci, hci)
amx_1	5724	1883	8779 (8297, 9298)	1521	5353 (5031, 5704)	976	2658 (2476, 2861)
amx_2	4599	1229	5121 (4775, 550)	984	3328 (3075, 3611)	714	1818 (1665, 1995)
amx_3	4633	1049	3594 (3328, 3891)	823	2512 (2311, 2740)	580	1454 (1327, 1607)
amx_4	3397	528	1086 (992, 1198)	359	532 (481, 604)	258	353 (320, 405)
A1.1	6298	527	917 (846, 1005)	332	474 (430, 537)	232	319 (287, 370)
A1.2	1791	315	628 (562, 713)	209	428 (372, 502)	151	275 (237, 392)
A5.1	14539	670	1296 (1198, 1414)	386	785 (707, 882)	226	422 (372, 489)
A5.2	6812	489	970 (885, 1073)	286	515 (461, 586)	173	238 (211, 283)
C1.1	7800	718	1527 (1412, 1659)	474	960 (875, 1064)	313	597 (520, 655)
C1.2	5766	661	1404 (1297, 1530)	462	863 (789, 954)	318	573 (517, 645)

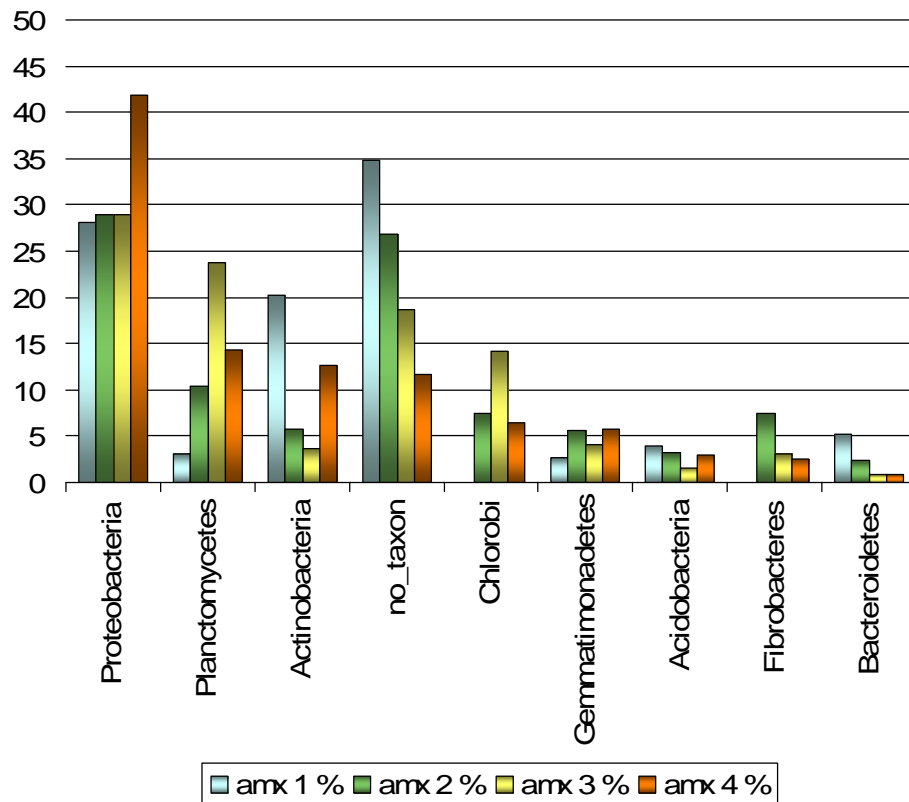
**Tab. 3.** ACE indexes for all samples. Lower confidence (lci) and higher confidence intervals are reported.

## 6.4 Sequences annotation and analysis of sample-related biodiversity

All sequences were compared using three databases of CAMERA and BLAST algorithm. For annotation, a “principle of coherence” (Par. 5.2) was followed, evaluating BLAST results and cluster analysis to obtain a single datum.

Considering the average similarity between sequences and the quality of information that would have resulted, the Prokaryotic Genomes database was referred to for biodiversity analysis.

For amx samples, the resulted bacterial community was assessed by specific types, their abundances and possible relationships (Fig. 16).



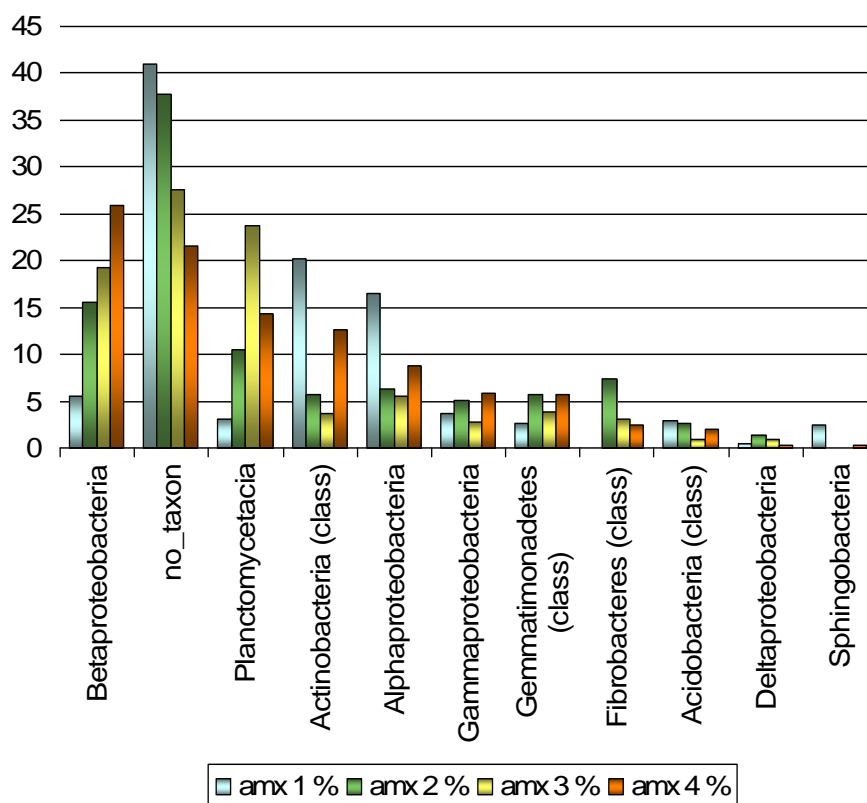
**Fig. 16.** Histogram displaying percentages of different phyla identified in the amx samples.

Proteobacteria and Actinobacteria represent the main constituents of the initial microbial consortium amx1 in which the other phyla were less represented.

Planctomycetes and Chlorobi growth may be favored by incubation conditions whereas Acidobacteria, Fibrobacters and Bacteroidetes were negatively selected. If percentage variations are considered, some correlations could be defined: a

decreasing number of Planctomycetes and Chlorobi coupled with an increasing number of Proteobacteria and Actinobacteria can be observed for the last amx4 sample.

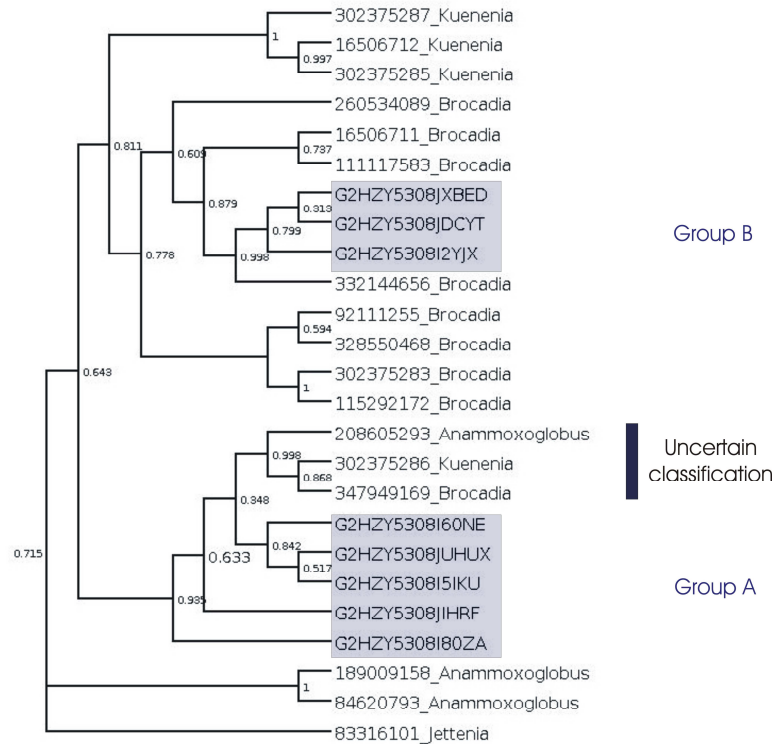
Considering the Eurotec WTT samples, growth conditions seem to have affected specific classes (Fig. 17).



**Fig. 17.** Histogram displaying percentages of different classes identified in the amx samples.

Within Proteobacteria, an increasing number of Betaproteobacteria was observed while Alphaproteobacteria decrease and Gammaproteobacteria remain unvaried. A positive experimental conditions-related effect was observed for Planctomycetes growth: a maximum value was detected for amx3 in which 23% of all the classes were represented by putative Anammox organisms.

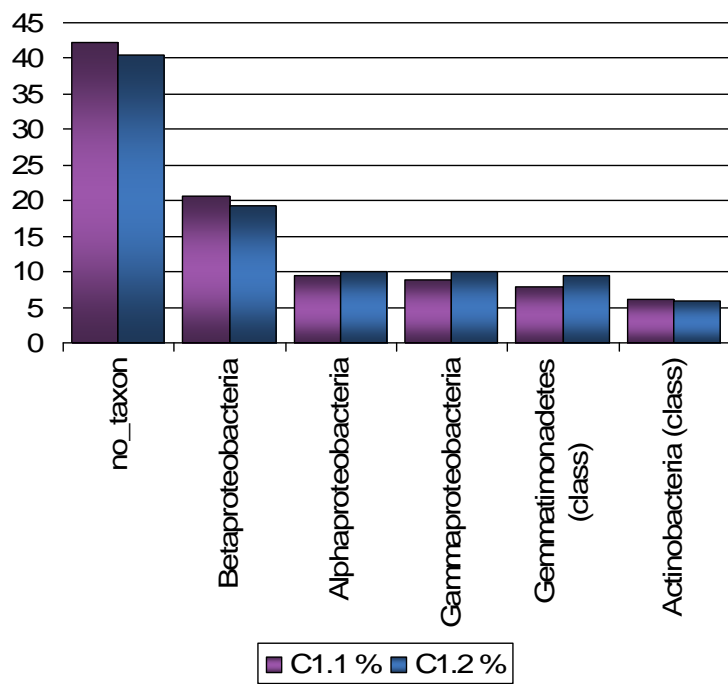
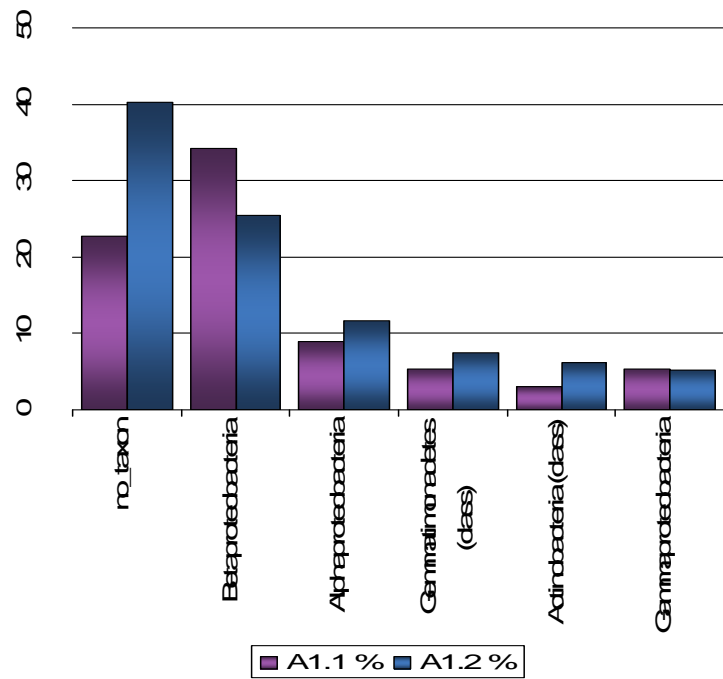
Sequences identified as Planctomycetes in sample amx4 were extracted and a phylogenetic tree was carried out considering other 16S rRNA Anammox-related sequences. Two different groups of bacteria were identified, related to different Anammox taxa (Fig. 18.).



**Fig. 18.** Neighbor-Joining phylogenetic tree. Anammox 16S rRNA sequences were downloaded by GenBank. Queries were annotated by a number (G.I., Accession Number) and a name that identifies the ‘candidate’ species. For amx4 sample, the representative sequences of each cluster at 97% of similarity were analyzed. Group A of amx4 sequences is sistergroup of an heterogeneous clade, while Group B is clearly related to *Brocadia spp.*

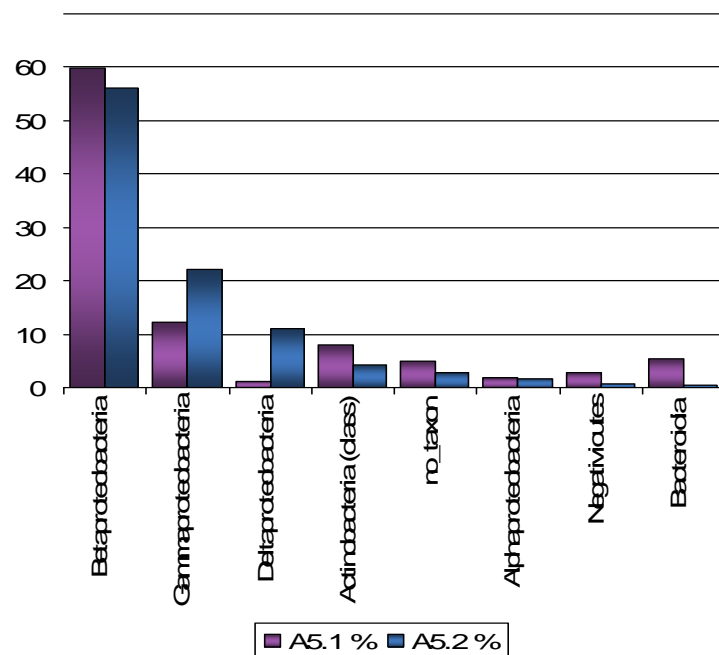
A decreasing number of not-annotated sequences was observed during time. Probably, the higher biodiversity recognized in the first sample involves several unknown organisms that are selected by the growth condition.

In the Milan bioreactors, the situation was rather different. A lower number of phyla was identified in each sample: a large amount of Betaproteobacteria can be observed in all samples and different trends were defined for the other Proteobacterial classes.



**Fig. 19.** Histograms displaying percentages of different classes identified in A1 and C1 batches, two timepoints are shown

The Proteobacteria population was always observed in sample A1, with comparable and stable values for Alpha and Gammaproteobacteria while in sample A5, an increasing number of Delta and Gammaproteobacteria is accompanied by a lower percentage of Alphaproteobacteria. A1 and C1 communities were comparable (Fig. 19,20).



**Fig. 20.** Histogram displaying percentages of different classes identified in A5 batches, two timepoints are shown.

A specific bacterial community was associated to each bioreactor and, in particular, to growth conditions: as defined by the percentage values, a selection was observed for samples derived by the same bioreactor but collected at different time points (summarized in Tab. 4).

ORDER	amx 1 %	amx 2 %	amx 3 %	amx 4 %	A1.1 %	A1.2 %	A5.1 %	A5.2 %	C1.1 %	C1.2 %
Acidimicrobiales	0,86	0,07	0,11	0,47	0,03	0,00	0,00	0,00	0,09	0,09
Acidobacteriales	2,94	2,57	0,91	1,94	0,19	0,11	0,00	0,00	0,19	0,78
Actinomycetales	10,82	2,66	1,79	1,92	0,99	2,74	7,73	4,08	2,89	2,90
Bacteroidales	0,00	0,00	0,00	0,00	0,19	0,00	5,31	0,57	0,19	0,23
Burkholderiales	1,92	7,07	5,13	10,87	11,75	12,68	57,28	54,31	9,44	8,76
Caulobacterales	0,58	0,20	1,83	3,68	0,33	1,51	0,00	0,06	2,84	0,40
Chromatiales	0,09	0,09	0,17	2,18	0,00	0,00	0,00	0,00	0,04	0,00
Clostridiales	0,23	0,28	0,11	0,06	0,14	0,34	1,16	0,15	0,65	0,56
Desulfovibrionales	0,00	0,00	0,00	0,00	0,00	0,06	1,09	11,07	0,00	0,00
Enterobacteriales	0,02	0,00	0,02	0,03	0,05	0,06	1,90	4,46	0,01	0,00
Fibrobacterales	0,00	7,40	3,04	2,47	0,00	0,00	0,00	0,00	0,00	0,00
Flavobacteriales	0,38	0,09	0,02	0,09	0,65	1,12	1,72	0,18	0,42	0,43
Gemmatimonadales	2,64	5,66	3,93	5,72	5,19	7,54	0,29	0,06	7,92	9,40
Hydrogenophilales	0,07	0,09	0,06	0,00	0,00	0,00	0,00	0,00	0,03	0,07
Ignavibacteriales	0,00	0,00	8,02	0,00	17,00	0,00	0,00	0,00	0,00	0,00
Neisseriales	0,03	0,20	0,19	0,41	1,05	1,12	1,67	0,47	0,71	0,85
Nitrosomonadales	0,09	0,07	0,11	0,12	21,02	11,34	0,00	0,03	9,67	8,55
Nitrospirales	0,05	0,04	0,02	0,03	0,00	0,00	0,00	0,00	0,00	0,00
Parvularculales	0,02	1,92	0,80	1,36	0,00	0,00	0,00	0,00	0,00	0,03
Planctomycetales	3,02	10,40	23,71	14,26	0,13	0,06	0,00	0,00	0,47	0,43
Pseudomonadales	0,30	0,30	0,09	0,38	1,23	1,40	8,45	16,36	1,77	2,24
Rhizobiales	4,63	1,57	1,57	2,39	7,05	8,99	1,74	1,40	5,49	5,60
Rhodobacterales	9,33	0,98	0,63	0,62	0,46	0,39	0,08	0,12	0,58	0,42
Rhodocyclales	0,30	2,37	5,07	4,12	0,11	0,06	0,08	0,18	0,19	0,49
Rubrobacterales	0,28	0,30	0,28	3,86	0,78	0,95	0,09	0,13	0,56	1,16
Selenomonadales	0,00	0,00	0,00	0,00	0,06	0,00	2,79	0,62	0,01	0,09
Solirubrobacterales	3,44	1,07	0,56	3,24	0,30	0,78	0,01	0,04	0,82	0,50
Sphingobacteriales	2,22	0,02	0,04	0,18	0,80	0,84	0,01	0,01	0,24	0,36
Vibrionales	0,00	0,00	0,00	0,00	0,02	0,00	1,33	1,13	0,00	0,00
Xanthomonadales	2,22	4,33	1,75	3,03	3,17	3,58	0,29	0,07	6,85	7,25
no_taxon	52,51	48,88	38,89	35,86	25,65	42,51	6,21	4,14	45,50	46,02
	> 2	> 5	> 10	> 20	> 30	> 40	> 50			

**Tab. 4.** List of the percentages of sequences attributed to each order for each sample (> 1% in at least one sample shown). Different colours show the highest percentages.

Databases that collect biological, biochemical and metabolic data allowing a comparison between community and explaining why two organisms can be found together, do not exist yet. A keywords-based search which may be useful to find a relation between samples was designed envisaging the biological role of a specific taxon (Squartini, 2011).

Some keywords related to the type of sample, were searched against the nucleotide GenBank database. The number of records obtained for each keyword considered alone, was compared with the number of records obtained if a phylum was considered as ‘secondary word’. Due to this strategy, phyla which ‘enrich’ specific keywords was related to the specific reaction, environment or characteristic.

Keyword searched were: anammox, nitrification, denitrification, wastewater, anaerobic. Histograms representing the most frequently phyla for each keyword are showed in Fig. 21.a and b.

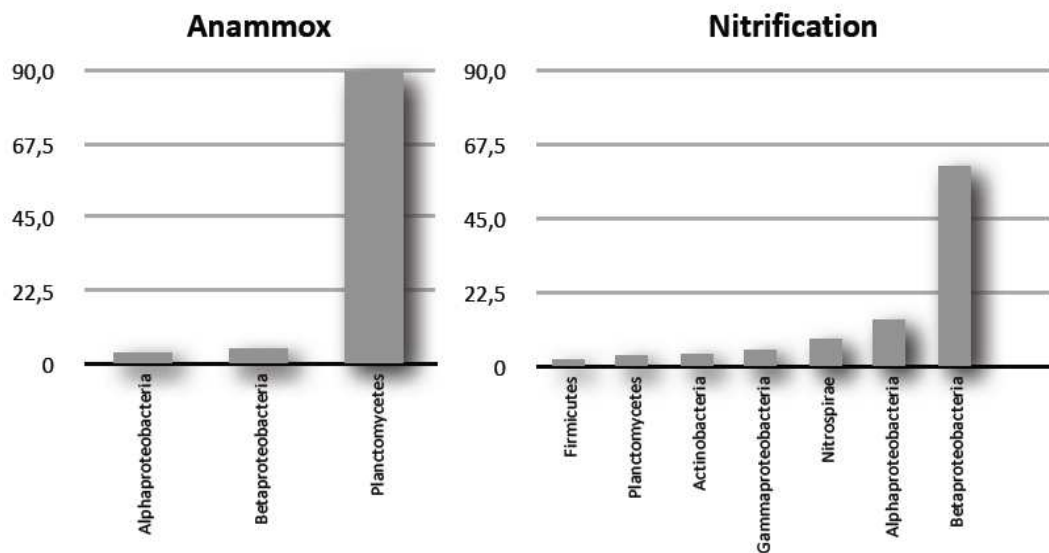
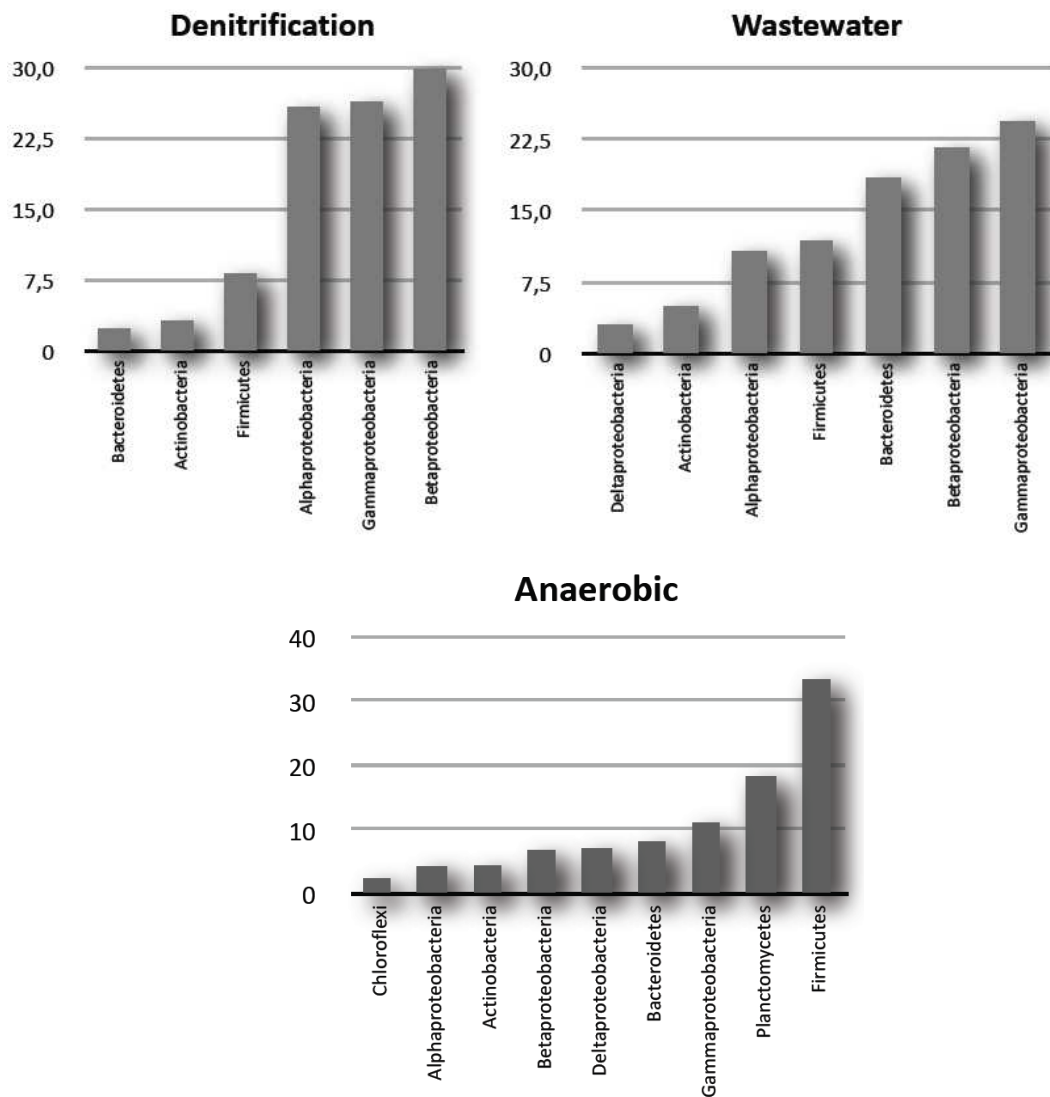


Fig. 21.a.



**Fig. 21.b.**

**Fig. 21.b. (21. ba)** Histograms representing the most frequently recurring phyla for each keyword searched in GenBank.

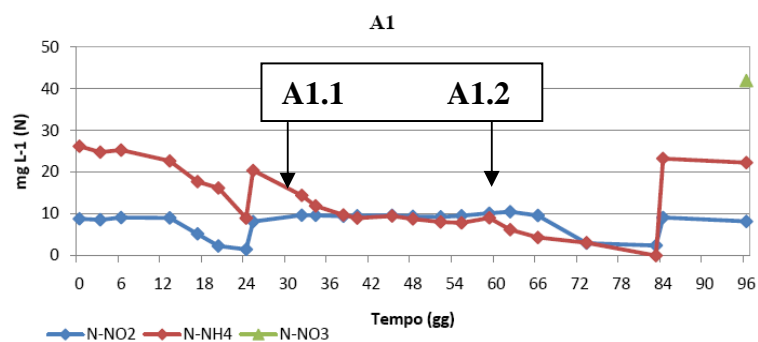
Analyzing these graphs, Planctomycetes are certainly the most representative phylum of the anammox reaction.

A close association between Betaproteobacteria and nitrification was reported and Betaproteobacteria, Gammaproteobacteria and Alphaproteobacteria are associated with denitrification. Moreover, Gamma and Betaproteobacteria and Bacteroidetes dominate wastewater environment while Firmicutes and Planctomycetes were the most abundant phyla if 'anaerobic' was used as keyword.

## 6.5 Genomic and biochemical data comparison

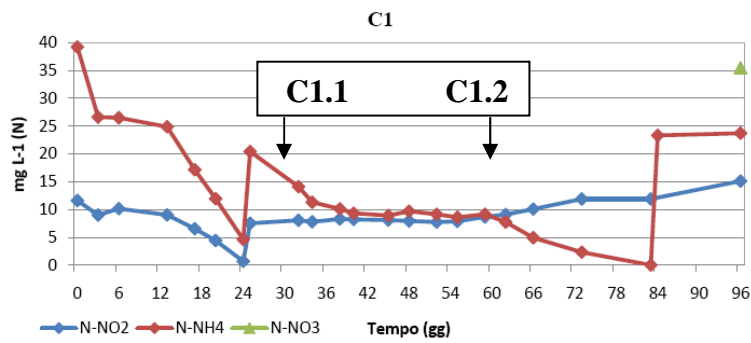
A community represented by Planctomycetes-like bacteria and Betaproteobacteria was developed starting from a denitrifying-nitrifying consortium by EurotecWTT. Nitrifiers and Anammox cooperate in ammonium oxidation as previously reported representing the most stable community able to maintain a functional Anammox-type microbial consortium.

Biochemical data were obtained by Prof. Adani's group analyzing nitrite and ammonia consumption: considering the reaction stoichiometry, a combined decreasing amount of such molecules, may indicate an Anammox-related activity. Days corresponding to sequencing were 30 and 60.



**Fig. 22.** Analysis of ammonia and nitrite consumption. A1.1 and A5.1 samples correspond to 30 and 60 days after the batch bioreactor start (courtesy of Prof Adani's group).

$\text{NO}_2^-$  and  $\text{NH}_4^+$  were equally consumed for A1 sample (Fig. 22): an anammox activity may be suggested. Results of 16S rRNA genes sequencing, in which a population of Proteobacteria (represented especially by Betaproteobacteria) and Actinobacteria may suggest an active nitrifying-denitrifying community (font: <http://biocyc.org>).

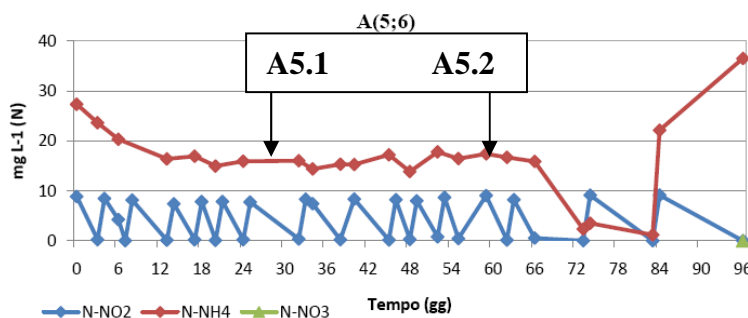


**Fig. 23.** Analysis of ammonia and nitrite consumption. C1.1 and C5.1 samples corresponding to 30 and 60 days after the batch bioreactor start (courtesy of Prof. Adani's group).

In sample C1, an increased COD/N=2 rate was tested. During time, a trend similar to that of the A1 sample was observed (Fig. 23).

Several species were identified only in A1 and C1: in particular, an increased number of Nitrosomonadales, Rhizobiales and Xanthomonadales. Focusing on Xanthomonadales, they were detected in all amx samples, A1 and C1 but an increased amount of such bacteria were observed only for sample C1.

It was remarkable that a small amount of Planctomycetes-related bacteria was found in A1 and C1 samples. The development of an Anammox community was allowed by the same environmental conditions that promoted a nitrifying-denitrifying consortium but an overall growth of this second population was probably favoured by such experimental conditions.



**Fig. 24.** Analysis of ammonia and nitrite consumption. A1.1 and A5.1 samples corresponding to 30 and 60 days after the batch bioreactor start. Two replicates were carried out for each bioreactor: figures show the average value evaluated between the two replicates A5 and A6 (courtesy of Prof Adani's group).

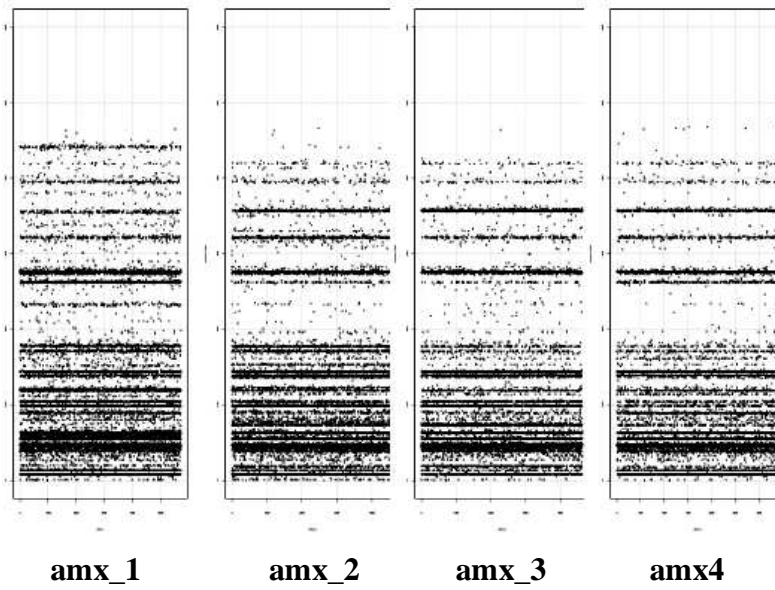
High rates of  $\text{NO}_2^-$  consumption were measured for A5 sample (Figure 18). This sample required daily nitrite supplementation, and a comparable trend was observed for a bioreactor in which denitrifying bacteria were used as inoculum. Several species were exclusively identified in this sample: the effects of the measured denitrification could be explained by an increasing percentage of Pseudomonadales, Actinomycetales, Desulfomonadales and Bacteroidales. Nitrosomonadales and Gemmatimonadales have not been identified, confirming the absence of nitrifying microorganisms annotated in a nitrification-denitrification communities evaluated for sample A1 and C1.

## 6.6 Virtual fingerprinting

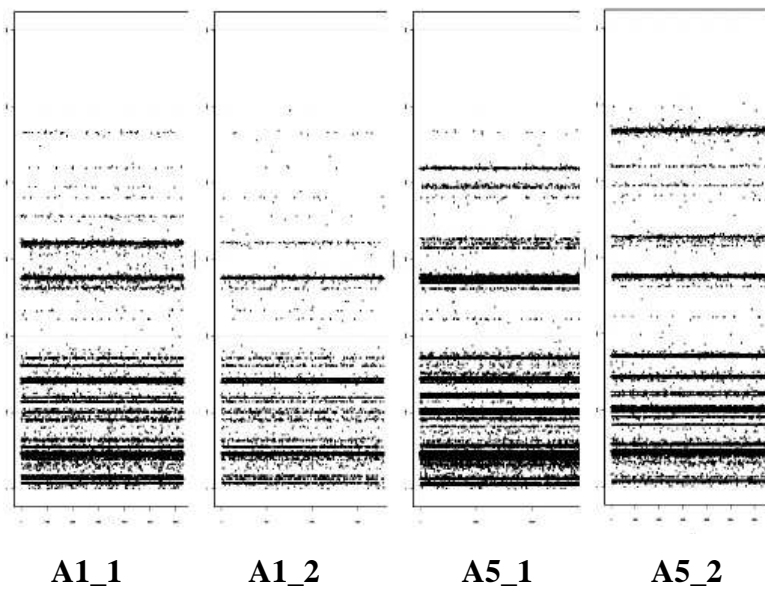
RFLP represents a useful procedure to seize bacterial communities. Variations of a restriction pattern may indicate variations in the microbial assemblage, representing a relevant information especially when a fully functional bioreactor must be maintained.

Based on this principle, a tool for virtual restriction reaction was developed to compare a standard and defined pattern with bands obtained by RFLP.

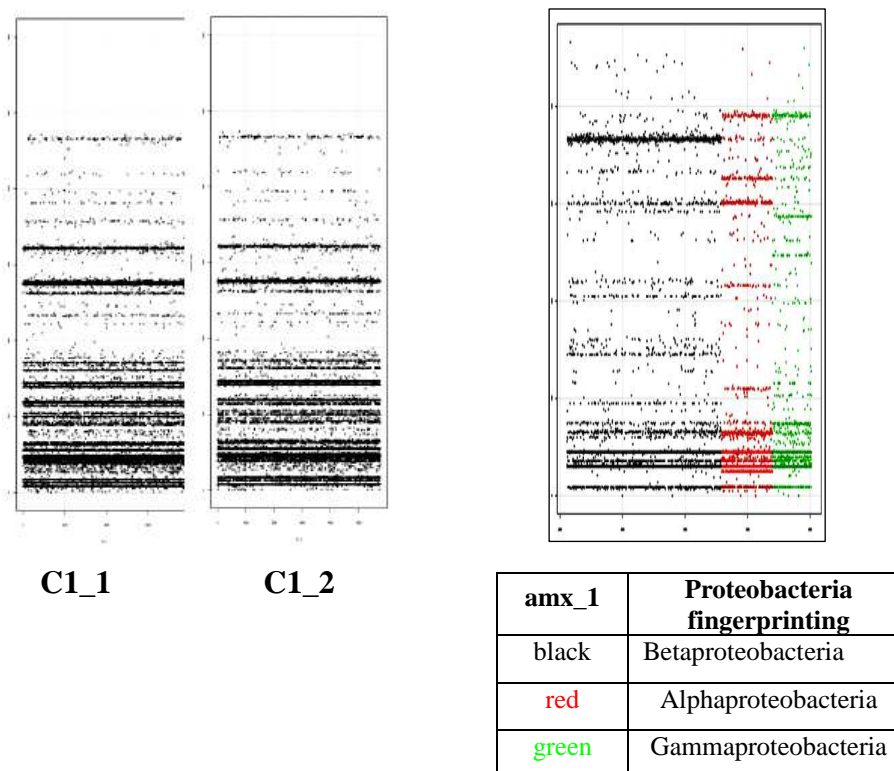
This tool was applied to all the 454 sequences. Using a MySQL databases in which restriction enzymes and recognition sites were collected, restriction sites for *HaeIII* and *Sau3a* were searched in all sequences and each sample, producing an ordered list representing the virtually digested fragments. A scatterplot was obtained using the R program and the result was visualized as a virtual gel electrophoresis (Fig. 25, 26).



**Fig. 25.** Scatterplot of virtual RFLP on amx 454 sequences using *Sau3A* and *HaeIII*.



**Fig. 26.** Scatterplot of virtual FRLP on Milan bioreactors



**Fig. 27.** Scatterplot of virtual RFLP on Milan bioreactor samples using *Sau3A* and *HaeIII*. On the right a virtual fingerprinting obtained by Proteobacteria identified sequences in amx\_1 sample.

Confirming the results of 'real' RFLP and sequencing annotation, the obtained pattern is specific for sample and time points considered. The higher heterogeneity was observed for amx1, representing the starting inoculum and confirming the other analyzed data.

A specific pattern may be defined considering all the amx samples in which an Anammox activity was measured. A different pattern characterizes the first and second time points of A1 and A5 samples supporting the variations in Betaproteobacteria and Alphaproteobacteria just described by annotation analysis (Fig. 19, 20). Moreover, a lower variation could be observed for sample C1 in which a stable community was observed for the first and the second time point sampling.

To find a correlation between bands and bacterial taxa, virtual restriction and annotation data can be considered together: sequences of sample amx4 belonging to Alpha, Beta and Gammaproteobacteria class were extracted and virtually cleaved using the *HaeIII* restriction site and revealing the specific taxon-related

pattern.

A specific Planctomycetes restriction map was defined for amx samples in which an Anammox activity was measured. A similar analysis was reported for Proteobacteria in Figure 20,

As the comparison between a scatterplot and an agarose electrophoresis gel is difficult due to the short lengths (less than 500 bp) and the low resolution power of agarose gel to separate short fragments, for such reason acrylamide gel or capillary electrophoresis are recommended to obtain a more informative resolution.

This analysis could represent a powerful method to monitoring a bioreactor and can define a 'standard pattern' in which every species is recognizable.

Once obtained a standard virtual fingerprinting for a microbial unknown community, RFLP can be used to monitor such population with respect to the biological activity of a bioreactor.

If a malfunctioning occurs, RFLP results can be compared with the standard pattern to test if a community corruption is taking place suggesting a possible way to solve the problem.

The development of anammox bioreactors is an ongoing task: the definition of an anammox-community related pattern, may be helpful for the evaluation of new starting inocula in which such community can exhibit a pattern closer to that of the functional bioreactor amx4 with respect to the undifferentiated amx1 and to the less performing amx2 and amx3.

## **6.7 Conclusions**

The development of an Anammox-type consortium was monitored in several bioreactors. The morphology of Anammox bacteria was investigated by electron microscopy while biochemical measurements certified the presence of an Anammox activity.

Specific microbial populations were characterized using semiquantitative PCR and RFLP: as expected, several microbial consortia were selected by different growth conditions and several metabolic capabilities were analyzed for each tested condition.

The presence of Anammox active bacteria was confirmed by the increasing

number of queries related to Planctomycetes identified due to high-throughput sequencing.

Considering the global involvement in nitrogen-compounds metabolism that characterizes a huge number of annotated species, a cooperative interaction could be suggested. Anammox bacteria were always detected in microbial communities where a coexistence between nitrifying and denitrifying bacteria was observed.

In the Eurotec WTT bioreactor (amx series), a functional Anammox community was selected while in the Milan bioreactors A1 and C1, a nitrification-denitrification equilibrium could be promoted comparing sequencing and biochemical data.

A competition for ammonium involving Anammox and nitrifying bacteria could be explained observing the amount of Planctomycetes bacteria in amx samples and *Nitrosomonas* in A1 and C1. A selection of denitrifying bacteria was determined by their higher replication rate: presence of this species as better competitor for nitrogen compounds determined a reduction of Anammox bacteria limiting them to a small subpopulation.

In the A5 bioreactor a denitrifying consortium was selected. Nitrate and nitrite consumption and ammonium accumulation could be explained considering nitrate and nitrite oxidative capabilities of such bacteria and the lack of any ammonium oxidation pathway. Specific community features were underlined by the high percentages of identified Burkholderiales, Pseudomonadales and competitors like Desulfurovibrionales.

Considering Eurotec WTT growth conditions applied as better conditions for the development of an Anammox type microbiota, other environments can be defined and implemented for reaction optimization. A specific pattern that defines such community was obtained due to a virtual fingerprinting. Due to this simple approach, new starter inocula could be assessed and followed through the enrichment of the Anammox community pattern.

Whole metagenomic and metatranscriptomic sequencing projects can be further devised to deeply analyze microbial community that can occur in an Anammox bioreactor and for a better understanding of the established relations that characterize these particular environments.

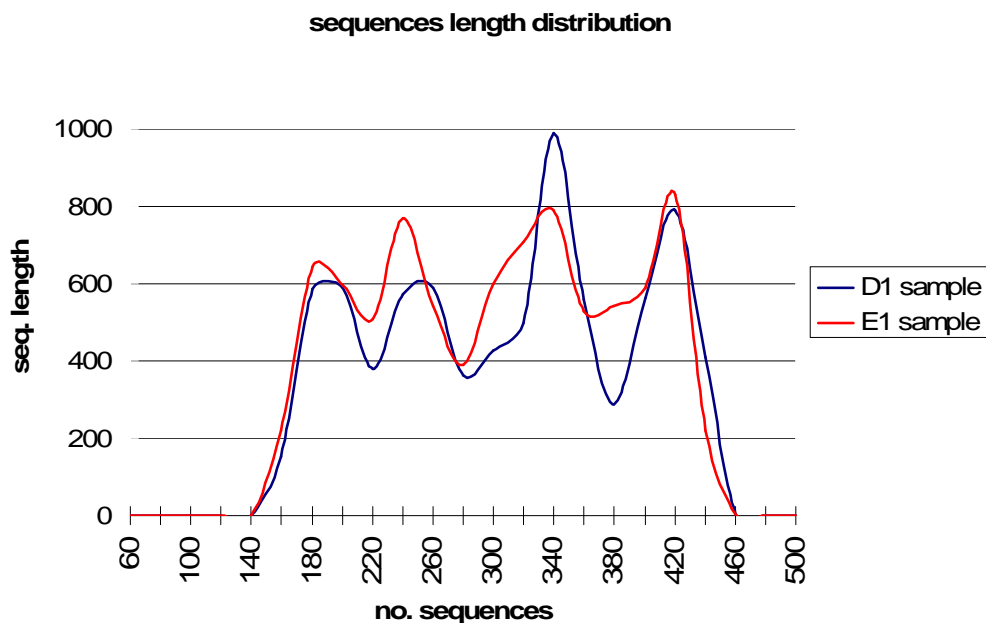
## Chapter 7. The dromedary rumen associated microflora

### 7.1 Sequencing results and data analysis

A total number of 23374 sequences was obtained, for a total of 6017298 bases.

Sequences were trimmed on the basis of length (>150 b) and quality: a length distribution analysis was performed to obtain information on quality results (Fig. 1).

For about 2480 reads, specific primer used for amplification was not tracked and these sequences were trimmed and 16213 sequences were analyzed. 7746 reads for D1 sample and 8467 for E1. A comparison with the three CAMERA databases was carried out and cluster analysis was implemented for sequence identification. Different results were obtained and for each dataset a different amount of not annotated sequences was obtained.



**Fig. 1.** Sequences length distribution. If compared with the Figure xxx, pag. xxx, a different sequencing performance can be observed considering Anammox and dromedary samples.

### 7.1.1 Cluster analysis

A cluster analysis was carried out on the basis of three similarity values, 100%, 99% and 97.5%.

Considering clusters at 100% of similarity, the number of unique sequences can be obtained, giving the first estimation about the sample-related heterogeneity.

Using 99% as threshold, sequencing errors and microvariability within organisms belonging to the same species (subpopulation or strains) were considered while a 97.5% of similarity cutoff value defined the putative number of species. An higher number of clusters were detected for atriplex-fed samples E1 with respect to hay-fed sample D1. Moreover, no significant variations were showed by 99% and 97.5% thresholds. Results are displayed in Fig 2.a and 2.b

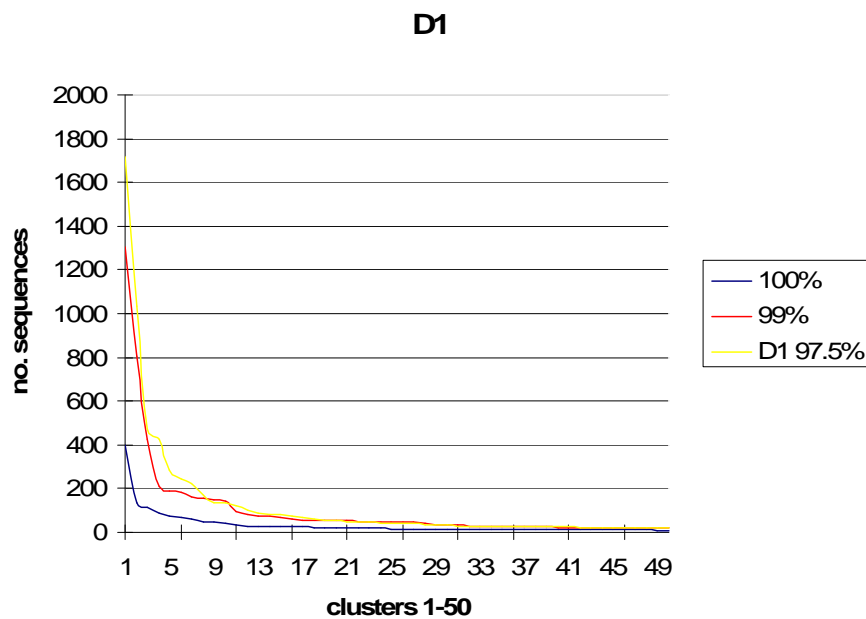
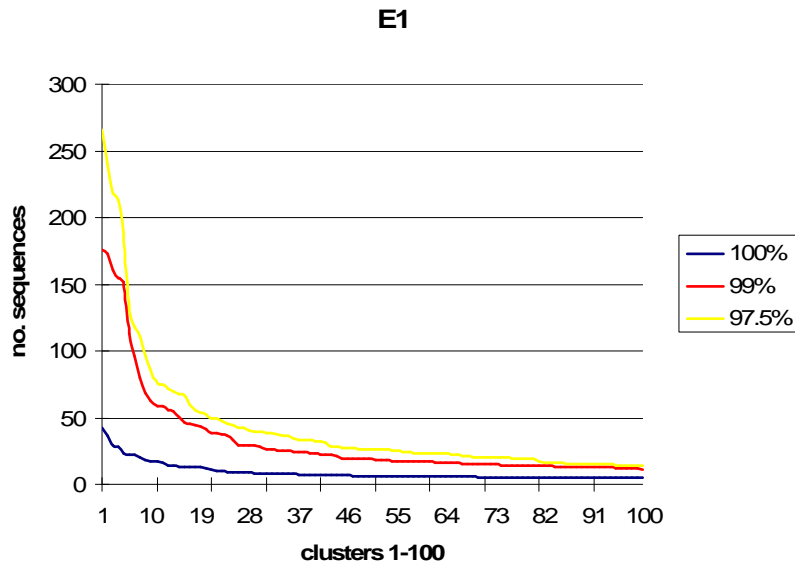


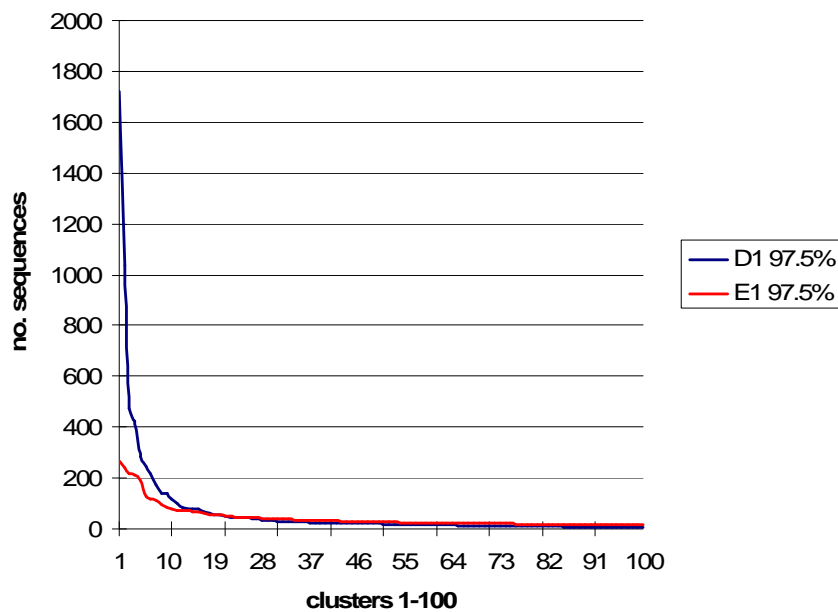
Fig. 2.a.



**Fig. 2.b.**

**Fig. 2.a., 2.b.** Number of reads clustered at 100%, 99% and 97.5% as similarity threshold. The 50 most representative clusters for sample D1 (Fig. 2.a) and the 100 for sample E1 (Fig. 2.b) are reported.

A results comparison considering 97.5% as threshold, showed an higher number of sequences grouped in a lower number of clusters for sample D1, indicating a low biodiversity for this sample (Fig. 3).



**Fig. 3.** Comparison between the number of sequences collected in D1 and E1 clusters (97%)

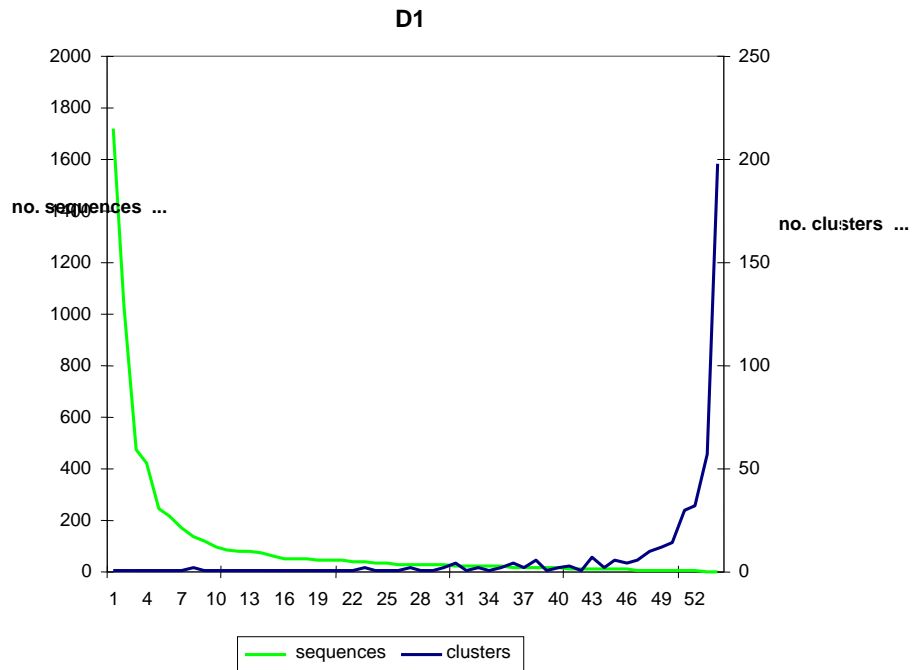


Fig. 4.a.

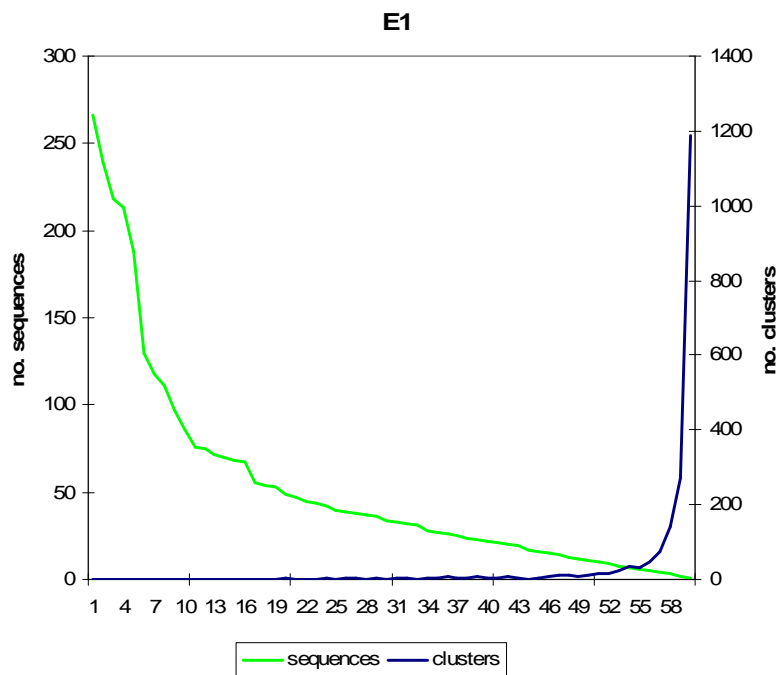


Fig. 4.b.

**Fig- 4.a., 4.b.** Comparison between the number of clusters and the number of sequences (not shown if < 10 sequences within a cluster). A higher number of sequences collected in a higher number of clusters can be evaluated for E1 by the lines interpolation (Fig 4.b).

The sequence number collected in each cluster was analyzed as cluster respective abundance index: for sample D1 an higher number of sequences were grouped in few clusters while in sample E1 a more homogeneous result was be observed.

E1 sequences are distributed among an higher number of clusters: several groups were represented by 20-70 sequences, confirming the higher heterogeneity of this sample (Fig. 4a, 4b).

## 7.2 Sequences annotation

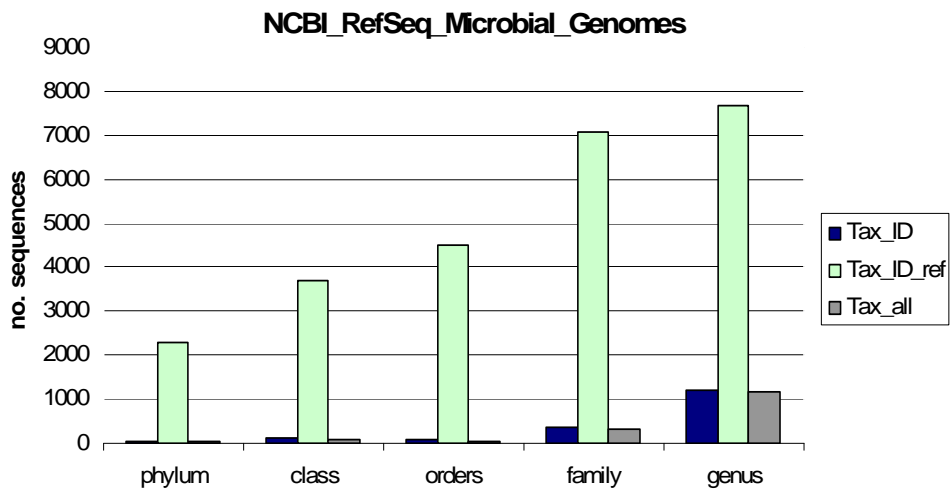


Fig. 5.a.

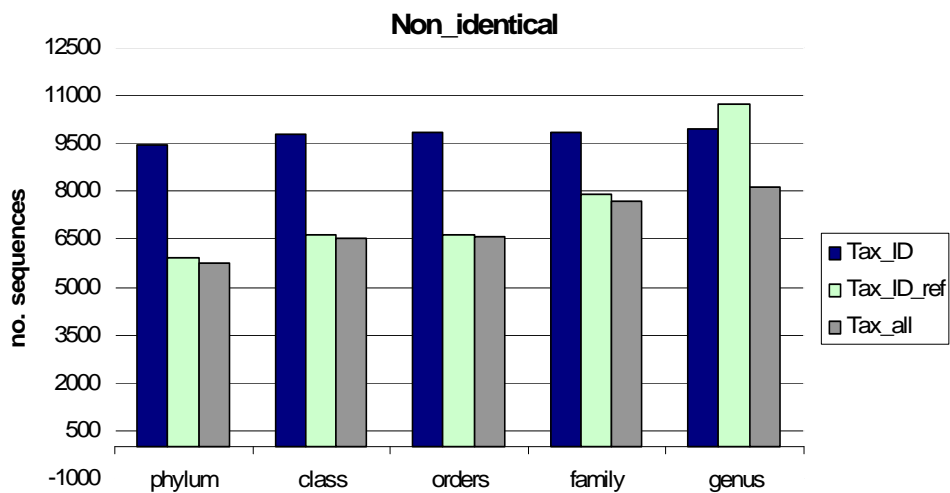


Fig. 5.b.

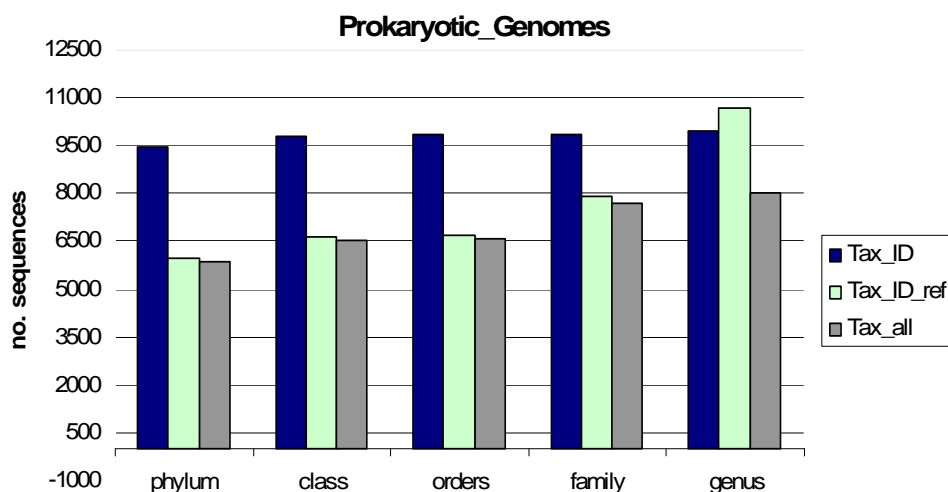


Fig. 5.c.

**Fig. 5.** Number of unidentified sequences considering the three databases of CAMERA. Tax\_ID: primary annotation; Tax\_ID\_Ref: annotation following the 'principle of coherence' described in chpt. 5 Tax\_all: identification considering primary annotation and annotation due to the 'principle of coherence'.

An opposite situation resulted by the application of the 'principle of coherence' described in par. 5.2 as protocol for queries annotation: information quality appeared to be closely related to the reference dataset. The annotation was difficultly interpretable and the wide sequences length distribution that characterized this 454-run may affect a cluster-based analysis in which a similarity evaluation between sequences was considered.

Considering NCBI\_RefSeq as reference database, an higher number of unidentified sequences ('no\_taxon') was identified following the principle of coherence (Fig. 5.a, Tax\_ID\_ref data). In agreement with the annotation protocol, a similar result may be explained analyzing the different annotation that characterize sequences belonging to a same cluster: differently annotated bacteria were grouped together. Surprisingly, a low number of unidentified hits were detected and decreased again if both cluster analysis and primary annotation data were considered.

Considering Non\_identical and Prokaryotic\_genomes databases, annotation results were comparable with the results obtained for the Anammox samples: at least 40% of sequences remain unidentified at higher taxonomic levels (family and genus) as shown in Fig. 5.b and c.

A deep annotation disagreement emerged by the comparison of D1 and E1

samples: while the most abundant phylum was represented by Firmicutes in both samples, several differences could be observed for lower taxonomic levels. Interestingly, the biodiversity revealed by annotation of atriplex-fed E1 sample was not detected for hay-fed D1 sample. Different bacterial classes, orders and families were identified but the estimated abundance was largely different. To clarify any mis-identification determined by the specific reference databases, a comparison of the results was carried out (Tab. 1).

<b>D1 sample</b>	<b>NCBI-Non Id %</b>	<b>NCBI-Prok Gen %</b>	<b>Non Id-Prok Gen %</b>	<b>All %</b>
<b>phylum</b>	0,06	0,18	24	75
<b>class</b>	0,03	0,03	24	74
<b>order</b>	0,03	0,07	32	66
<b>family</b>	0,02	0,06	35	64
<b>genus</b>	0,03	0,06	65	34
<b>species</b>	0,01	0,02	68	0,07

<b>E1 sample</b>	<b>NCBI-Non Id %</b>	<b>NCBI-Prok Gen %</b>	<b>Non Id-Prok Gen %</b>	<b>All %</b>
<b>phylum</b>	0,6	0,2	89	7
<b>class</b>	0,28	0,1	91	5
<b>order</b>	0,23	0,1	92	4
<b>family</b>	0,23	0,05	90	6
<b>genus</b>	0,21	0,05	81	14
<b>species</b>	0,5	0,05	90	0,5

**Tab. 1.** Percentages of sequences giving the same annotation (identified by the same Tax\_ID number). NCBI: NCBI\_RefSeq\_microbial\_Genomes; Non Id: Non\_identical;Prok Gen: Prokaryotic\_genomes; All: sequences shared between the three databases.

On the basis of Taxon ID (which officially identifies every taxonomic lineage), the percentage of sequences giving the same annotation were reported. Few completely not-shared sequences were detected.

For the D1 sample, 75% of sequences were associated with a same Taxon ID by the three databases.

A remaining 24% were recognized belonging to the same phylum by the Non\_identical and Prokaryotic\_genomes databases giving a different annotation for NCBI\_RefSeq\_Microbial\_Genomes. An higher number of shared sequences was identified at lower taxonomic levels in this sample: this value is related with the increasing amount of sequences annotated as 'no\_taxon' and reaches the 50%

for Non\_identical and Prokaryotic\_Genomes.

The situation appeared different considering sample E1. The number of shared sequences between databases has not a 'regular distribution' as reported for D1 sample.

From phylum to genus, a decreasing number of hits have been annotated with the same Tax\_ID for the three databases and the higher number of sequences is shared by Non\_identical and Prokaryotic\_genomes. On one hand, the 'no\_taxon' number increases giving information about the databases' richness. On the other hand, this may be considered as a biodiversity-related effect in which an higher number of unknown organisms were detected in this specific sample.

This hypothesis was confirmed by a differentially-annotated sequences analysis: considering sample D1, Actinobacteria and Firmicutes identified by NCBI\_RefSeq\_Microbial\_Genomes are recognized as 'no\_taxon' by the other databases (~10%). Within these phyla, the same differences were also observed at lower taxonomic levels, for Actinomycetales and Bacilli class.

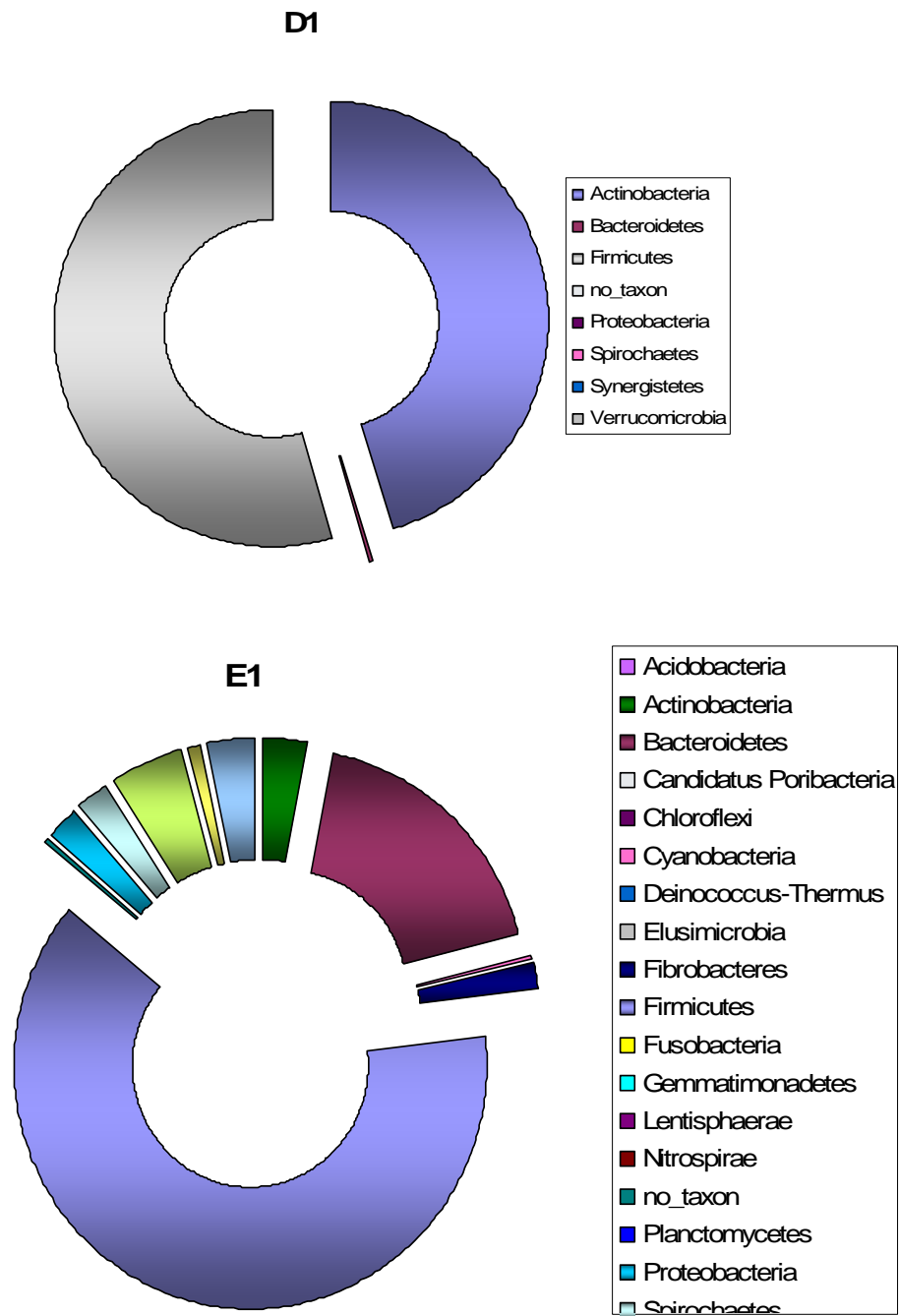
For sample E1 the same apparent mis-annotation can be observed. Several Firmicutes and Bacteroidetes identified using NCBI\_RefSeq, were annotated as 'no\_taxon' (~60%) by other databases and the same trend was observed for Clostridiales and Bacteroidales class.

A different information which characterizes each database was underlined by these comparisons.

The presence of a huge amount of data concerning shotgun sequencing and metagenomic analysis in NCBI\_RefSeq\_Microbial\_genomes evidently allows to obtain a more accurate information.

Without considering the not-assigned reads, identified species score are rather similar for all the databases and several differences were detected with respect to the specific sample.

### 7.3 Sequences annotation and analysis of sample-related biodiversity

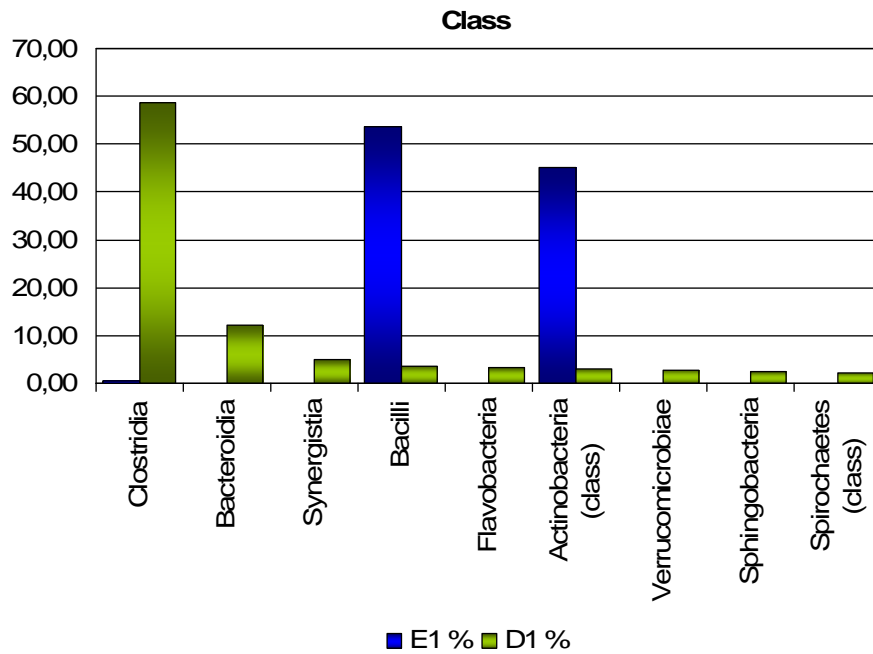


**Fig. 6.** Annotation results using NCBI\_RefSeq\_Microbial\_Genomes as reference database. An higher number of taxa characterizes sample E1. Two phyla represent the 99% of the entire biodiversity detected in sample D1.

As indicated by cluster analysis, an higher biodiversity was defined for Atriplex-fed camel. Firmicutes and Bacteroidetes, Actinobacteria and Synergistetes dominate the E1 bacterial community where Firmicutes and Actinobacteria

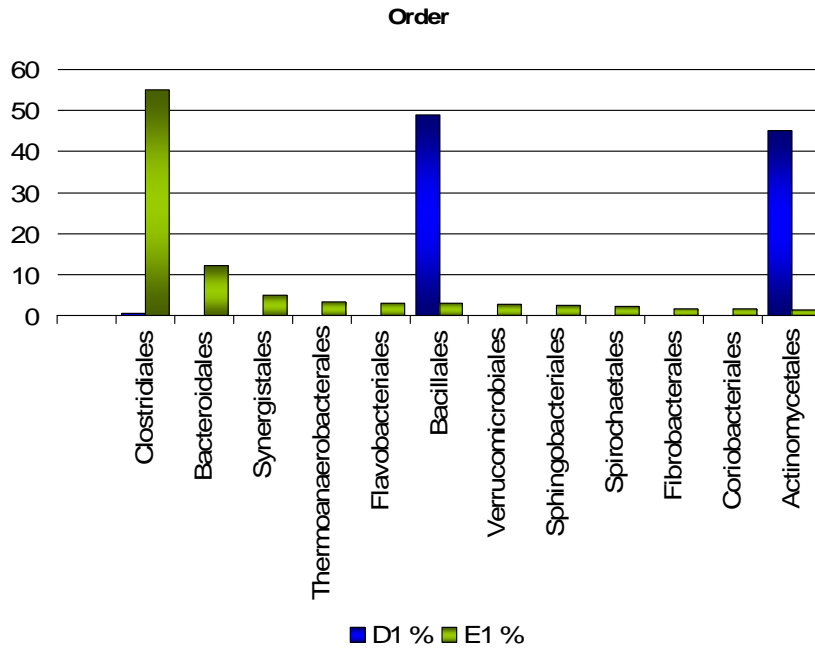
represented the most abundant phyla in the D1 sample while other phyla are poorly represented (Fig. 6).

While the most abundant phylum in dromedary rumen seems to be represented by Firmicutes, an interesting evidence was observed comparing samples at higher taxonomy levels.



**Fig. 7.** Taxa identified at class level. Bacilli represent the most abundant Firmicutes-related taxa identified in 1 sample while the Clostridia class dominates E1 sample.

At class level, several differences between the two samples were observed, in particular within the Firmicutes phylum (Fig. 7). While Bacilli represents the most abundant class recognized in the hay-fed D1 sample, in the atriplex-fed sample E1, Clostridia is the most abundant. This result was similar for the three databases considered and was confirmed by a comparison with the Classifier-tool output of the Ribosomal Database Project online source.



**Fig. 8.** Histogram displaying percentages of different orders identified in D1 and E1 samples.

A similar results was obtained considering the order as taxonomic level (Fig. 8).

At family level, Firmicutes population of E1 seems to be dominated by Clostridiaceae (16% of sequences) followed by Ruminococcaceae and Eubacteriaceae. 10% was constituted by Lachnospiraceae (in which the genus *Butyrivibrio* is classified) and 6% by Prevotellaceae.

The Bacteroidetes phylum, represents the other most abundant taxon identified in the E1 sample: 9% of the entire bacterial population is represented by Flavobacteriaceae, Bacteroidaceae and Porphyromonadaceae classes.

A completely different situation was observed for sample D1. While a limited number of clusters at high taxonomic levels was previously suggested by cluster analysis, at class level the predominance of Bacillaceae and Corynebacteriaceae was clearly supported. Whithin the Firmicutes phylum Bacilli, Lactobacillaceae and Planococcaceae represent the most abundant classes (Table 2).

FAMILY	D1 %	E1 %
Clostridiaceae	0,08	16,02
Ruminococcaceae	0,04	11,24
Eubacteriaceae	0,08	10,86
Lachnospiraceae	0,13	10,07
Prevotellaceae	0,03	6,33
Synergistaceae	0,03	5,00
no_taxon	0,09	3,88
Flavobacteriaceae	0,10	3,11
Bacteroidaceae	0,00	2,90
Porphyromonadaceae	0,05	2,82
Peptococcaceae	0,00	2,71
Sphingobacteriaceae	0,01	2,35
Verrucomicrobia subdivision 3	0,01	2,22
Spirochaetaceae	0,01	2,17
Paenibacillaceae	0,05	2,07
Thermoanaerobacterales Family III. Incertae Sedis	0,03	2,02
Fibrobacteraceae	0,00	1,68
Coriobacteriaceae	0,04	1,53
uncultured	0,49	1,35
Thermoanaerobacterales Family IV. Incertae Sedis	0,01	1,04
Acholeplasmataceae	0,00	0,77
Bacillaceae	43,91	0,61
Verrucomicrobiaceae	0,00	0,54
Moraxellaceae	0,00	0,49
Methylacidiphilaceae	0,00	0,40
Campylobacteraceae	0,00	0,36
Erysipelotrichaceae	0,01	0,34
Corynebacteriaceae	43,31	0,32
Dermacoccaceae	0,14	0,31
Peptostreptococcaceae	0,06	0,31
Syntrophomonadaceae	0,00	0,28
Beutenbergiaceae	0,77	0,25
Aerococcaceae	0,84	0,20
Staphylococcaceae	0,71	0,19
Micrococcaceae	0,42	0,18
Clostridiales Family XI. Incertae Sedis	0,01	0,15
Lactobacillaceae	3,11	0,15
Thermoanaerobacteraceae	0,00	0,15
Pasteurellaceae	0,00	0,14
Planococcaceae	4,16	0,14

**Table 2.** List of the percentages of sequences attributed to each family. Different colours shows the higher percentages.

#### 7.4 Microbiological considerations

A microbial community comparable with other ruminal consortia was found in atriplex-fed sample: due to their fermentative capabilities, anaerobic Gram positive bacteria are primary involved in fatty acids production.

Cellulosolytic activities were described for several *Clostridium* related organisms. Such species can synthesize a multi-enzymatic cellulosome (Bayer et al., 1998) which is directly involved in the cellulolysis process and are characterized by an high proteolytic activity. Butyrate and acetate represent the main fermentation products and are fundamental for the host glycogen synthesis pathway.

Other taxa previously described in several mammalian digestive tracts can be observed: *Eubacterium* seems to be a common dromedary rumen associated microorganism (Samsudin et al., 2011; Li et al., 2011) and was found in the digestive tract of other animals and in adult human distal gut (Barcenilla et al., 2000; Duncan et al., 2008; Mahowald et al., 2009). This species represents one of the major soluble sugars fermenter (Yoda et al., 2005) and fermentation products comprise formate, acetate and butyrate (Van Gylswyk et al., 1986).

*Prevotella* is one of the most abundant species isolated from the rumen and from the indgut of other mammalian (Avgustin et al., 1997). Species belonging to this genus are characterized by amilolytic, xylanolytic, pectinolytic activities and produce formate, acetate, propionate and butyrate as the major fermentation products.

*Sphingobacteria* and *Verrucomicrobia*, *Paenibacillus* and *Thermoanaerobacter* are characterized by xylanolytic activity (Hespell, 1992; Pason et al., 2006; Park et al., 2010; Wang et al. 2011; Zhou et al., 2010) and *Spirochaetaceae* are characterized by a pectinolytic activity (Wojciechowicz et al., 1979; Ziółcki et al., 1980).

While a cellulosolytic activity was reported for several identified bacteria in the E1 sample, *Porphyromonadaceae* and *Coriobacteriaceae* lack in this pathway. These bacterial families belonging to *Bacteroidetes* and *Actinobacteria* phyla are characterized by an high proteolytic activity and in other ruminants, were identified as components of the liquid phase (De Menezes et. al, 2011). They can

contribute to NH<sub>3</sub> accumulation which represent a problem in relation to the rumination process. A detoxification capability characterizes several Coriobacteria related species which can easily metabolize forages containing nitrotoxins (Anderson et al., 1996).

Due to the specific activities recognized in the identified species, an hypothesis may be formulated about the contribute of the bacterial taxa found in sample E1.

Due to hemycellulolytic activity, a role of early degrader may be hypotized for *Prevotella*-related organisms, supported by Sphingobacteria and other xylanolytic species. *Clostridium* and *Ruminococcus* may represent the primary degraders which can adhere to substrate and carry out the cellulosome-mediated cellulolysis. Produced free sugars may be metabolized by *Eubacterium* species. Other bacteria unable to hydrolyze cellulose, like Coriobacteriaceae were detected: some of these may increase the activity of cellulolytic bacteria (as reported for other rumen consortia) and may play a role in digestive process and food toxic-substances damage prevention.

A limited information was available observing hay-fed D1 sample whereas two phyla seems to represent the entire microbial consortium. Bacillaceae and Corynebacteriaceae were identified as the most abundant families and a similar situation was not reported for any previously analyzed rumen.

*Bacillus spp.* were detected in other ruminal samples as components of the liquid phase (Williams et al., 2008). A cellulolytic activity was detected in micro-aerophilic conditions (Fujimoto et al., 2011) and explained by the detection of endoglucanases (Bischoff et al., 2006) but its contribution in cellulolysis has not been defined. *Bacillus* species may be used as 'feed supplement' improving ruminal microbial balance (Fuller et al., 1989; Krehbiel et al., 2003; Qiao et al., 2008).

While other Actinobacteria were identified as ruminal associated species (An et al., 2006) poor relations between Corynebacteriaceae and ruminal processes were detected. Small communities of Corynebacteriaceae were detected in ruminal consortia (Leng et al., 2011) and several species are related to milk (Callon et al., 2007) and milk-derived products as cheese (Monnet et al., 2012; Mounier et al., 2007). Corynebacteriaceae were also identified as causative agents of infectious

diseases in ruminants (Loste et al., 2005).

If the other most abundant families identified in D1 sample are considered, for some Planococcaceae bacteria a cellulolytic activity has been reported (Choi et al., 2007; Fayyaz-ur-Rehman et al., 2009) and they were found in association with other ruminants (Yang et al., 2010; Kim et al., 2011).

*Lactobacillus* species are used as diet integrator (Mohammed et al., 2012) but were also reported as component of the ruminal 'normal flora'.

For the D1 sample, due to a limited number of identified taxa, it appears preliminary to formulate any hypothesis concerning exact biological roles for such organisms.

Several considerations may be put forward to explain this result. Sequencing run may be affected by polyclonal sequencing, in which several copies of the same template may be amplified and spread on emulsion PCR-beads determining an over-production of identical sequences. Other anomalies may be related to sample. Use of Bacillaceae as hay diet integrator may explain the amount of identified *Bacillus* species while Corynebacteriaceae suggests that some infectious diseases may be occurred increasing bacterial growth rate and the development of a cooperative interaction.

The rumen microbiota appears to be complex, species-rich in ways that were not predicted by classic knowledge on ruminants' physiology and symbioses.

The community structure appears to be strongly dependent on the diet followed by the animal. Further studies will be needed to corroborate these data and verify other effects.

## **Chapter 8. Conclusions**

The study has allowed to explore the problem of microbial taxonomy upon defining the adequate tools and cutoff values to operate with workable concepts. The theory has been tested on real-life microbial communities of natural and artificial kind. The large body of acquired data has been an ideal playground to devise and implement programs and specific scripts to handle the information and to draw patterns that start to emerge on the constraints imposed by the environment to microbial assemblages. The whole array of data gathered has enabled to step forward in the understanding the relative weights of constancy and variability in the structure of living communities. At the same time besides the molecular ecology lesson that data have taught us, an important series of clues have emerged that are of direct practical use in the monitoring and handling of cell-based reactors in the rewarding field of applied biotechnologies.

## Bibliography

- Achenbach L.A., Carey J., Madigan M.T. Photosynthetic and phylogenetic primers for detection of anoxygenic phototrophs in natural environments. 2001, *Applied and Environmental Microbiology*; 67: 2922-2926
- Ahn Y.H, Sustainable nitrogen elimination biotechnologies: A review. 2006, *Process Biochemistry*; 41:1709 – 1721.
- Allemand F., Mathy N., Brechemier-Baey D., Condon C. The 5S rRNA maturase, ribonuclease M5, is a Toprim domain family member. 2005, *Nucleic Acids Research.*; 33: 4368-76.
- Amman R.I., Lin C., Key R., Montgomery L., Stahl D.A. Diversity among Fibrobacter isolates: towards a phylogenetic classification. 1992, *Systematic and Applied Microbiology*, 15: 23-31
- Amman R.I, Ludwig W., Schleifer K.H. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. 1995, *Microbiol Rev.*; 59: 143-69.
- An D., Cai S., Dong X.. *Actinomyces ruminicola* sp. nov., isolated from cattle rumen. 2006, *International Journal of Systematic and Evolutionary Microbiology*; 56: 2043-8.
- Anderson R.C., Rasmussen M.A., Allison MJ. Enrichment and isolation of a nitropropanol-metabolizing bacterium from the rumen. 1996, *Applied and Environmental Microbiology*; 62: 3885-6.
- Arkov A.L., Freistroffer D.V., Ehrenberg M., Murgola E.J. Mutations in RNAs of both ribosomal subunits cause defects in translation termination. 1998, *EMBO Journal*; 2-17: 1507-14.
- Arp, D.J., Sayavedra-Soto L.A., Hommes N.G.. Molecular biology and biochemistry of ammonia oxidation by *Nitrosomonas europaea*. 2002, *Archives of Microbiology*; 178: 250–255.
- Attwood G.T., Reilly K., Patel B.K. *Clostridium proteoclasticum* sp. nov., a novel proteolytic bacterium from the bovine rumen. 1996, *International Journal of Systematic and Evolutionary Microbiology*; 46: 753-8.
- Avgustin G., Wallace R.J., Flint H.J. Phenotypic diversity among ruminal isolates of *Prevotella ruminicola*: proposal of *Prevotella brevis* sp. nov., *Prevotella bryantii* sp. nov., and *Prevotella albensis* sp. nov. and redefinition of *Prevotella ruminicola*. 1997, *International Journal of Systematic Bacteriology*; 47: 284-288.
- Bakin A., Ofengand J. Four newly located pseudouridylate residues in *Escherichia coli* 23S ribosomal RNA are all at the peptidyltransferase center: analysis by the application of a new sequencing technique. 1993, *Biochemistry*; 32: 9754-62.
- Balcazar J.L., de Blas U., Ruiz-Zarzuola I., Vendrell D., Girones O., Musquiz J.L. Sequencing of variable regions of the 16S rRNA gene for identification of lactic acid bacteria isolated from the intestinal microbiota of health salmonids. 2007, *Comparative Immunology, Microbiology & Infectious Diseases*; 30: 111-118.
- Balci S., Dincel Y. Ammonium ion absorption with sepolite: use of transient uptake method. 2002, *Chemical Engineering and Processing*; 41, 70-85.

- Bansal A.K., Meyer T.E. Evolutionary analysis by whole-genome comparison. 2002, *Journal of Bacteriology*, 184: 2260-2272.
- Barcenilla A, Pryde S.E., Martin J.C, Duncan S.H., Stewart C.S., Henderson C., Flint H.J. Phylogenetic relationships of butyrate-producing bacteria from the human gut. 2000, *Applied and Environmental Microbiology*; 66: 1654-1661.
- Baxter-Roshek J.L., Petrov A.N., Dinman J.D. Optimization of ribosome structure and function by rRNA base modification. 2007, *PLoS One*; 2: e174.
- Bayer E.A, Lamed R., White B.A, Flint H.J. From cellulosomes to cellulosomes. 2008, *Chemical Record.*; 8: 364-77.
- Berg-Miller M.E., Antonopoulos D.A., Rincon M.T., Band M., Bari A., Akraiko T., Hernandez A., Thimmapuram J, Henrissat B, Coutinho PM, Borovok I, Jindou S, Lamed R, Flint HJ, Bayer EA, White BA. Diversity and strain specificity of plant cell wall degrading enzymes revealed by the draft genome of *Ruminococcus flavefaciens* FD-1. 2009, *PLoS One.*; 4: e6650.
- Bergman E.N. Energy contributions of volatile fatty acids from the gastrointestinal tract in various species. 1990, *Physiological Reviews*; 70: 567-90.
- Bergmann, D.J., Hooper A.B. Cytochrome P460 of *Nitrosomonas europaea*. 2003, *The Journal of Biochemistry*; 270: 1935–1941 .
- Bischoff K.M., Rooney A.P., Li X.L., Liu S., Hughes S.R. Purification and characterization of a family 5 endoglucanase from a moderately thermophilic strain of *Bacillus licheniformis*. 2006, *Biotechnology Letters*; 28: 1761-5.
- Bosshard P.P., Abels S, Zbinden R., Böttger E.C., Altwegg M. Ribosomal DNA sequencing for identification of aerobic gram-positive rods in the clinical laboratory (an 18-month evaluation). 2003, *Journal of Clinical Microbiology*; 41: 4134-40.
- Bouraoui H., Vendramin, Squartini A, and Haddi ML Taxonomical analysis of the bacterial content of the camel rumen fluid. 2006, 4th International Conference on Biological Sciences, Tanta Egypt, 1-2 nov.
- Brink M.F., Verbeet M.P., de Boer H.A. Formation of the central pseudoknot in 16S rRNA is essential for initiation of translation. 1993, *EMBO Journal*; 12: 3987-96.
- Britton R.A., Wen T., Schaefer L., Pellegrini O., Uicker W.C., Mathy N., Tobin C., Daou R., Szyk J., Condon C. Maturation of the 5' end of *Bacillus subtilis* 16S rRNA by the essential ribonuclease YkqC/RNase J1. 2007, *Molecular Microbiology*; 63: 127-38.
- Brochier C., Philippe H. Phylogeny: a non-hyperthermophilic ancestor for bacteria. 2002, *Nature*; 417: 244.
- Browlee G.G., Sanger F., Barrel B.G. Nucleotide sequence of 5S-ribosomal RNA from *Escherichia coli*. 1967, *Nature*; 215: 735-736
- Brulc J.M., Antonopoulos D.A., Miller M.E., Wilson M.K., Yannarell A.C., Dinsdale E.A., Edwards R.E., Frank E.D., Emerson J.B., Wacklin P., Coutinho P.M, Henrissat B., Nelson K.E, White B.A. Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. 2009, *PNAS*; 106: 1948-53.
- Brumm P., Mead D., Boyum J., Drinkwater C., Deneke J., Gowda K., Stevenson D., Weimer P. Functional annotation of *Fibrobacter succinogenes* S85 carbohydrate active

enzymes. 2011, Applied Biochemistry and Biotechnology; 163: 649-57.

- Bryant M.P., Small N. The anaerobic monotrichous butyric acid-producing curved rod-shaped bacteria of the rumen. 1956, Journal of Bacteriology; 72: 16-21.
- Burton R.A., Gidley M.J., Fincher G.B. Heterogeneity in the chemistry, structure and function of plant cell walls. 2010, Nature Chemical Biology; 6: 724-32.
- Byrne N, Strous M, Crépeau V, Kartal B, Birrien JL, Schmid M, Lesongeur F, Schouten S, Jaeschke A, Jetten MSM, Prieur D, Godfroy A. Presence and activity of anaerobic ammonium-oxidizing bacteria at deep-sea hydrothermal vents. 2009, ISME Journal; 3: 117-123
- Callon C., Duthoit F., Delbès C., Ferrand M., Le Frileux Y., De Crémoux R., Montel M.C. Stability of microbial communities in goat milk during a lactation year: molecular approaches. 2007, Systematic and Applied Microbiology.; 30: 547-60.
- Case R.J., Boucher Y., Dahllöf I., Holmström C., Doolittle W.F., Kjelleberg S. Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. 2007, Applied and Environmental Microbiology; 73: 278-88.
- Chakravorty S., Helb D., Burday M., Connell N., Alland D. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. 2007, Journal of Microbiological Methods; 69: 330-9.
- Chao A. Nonparametric Estimation of the Number of Classes in a Population. 1984, Scandinavian Journal of Statistics 11: 265-270,
- Chao A. Stimating the population size for capture-recapture data with unequal catchability. 1987, Biometrics.; 43: 783-91.
- Chao A, Lee S.M. Estimating the number of classes via sample coverage. 1992, Journal of the American Statistical Association; 87: 210-217
- Chen H, Liu S, Yang F, Xue Y, Wang T. The development of simultaneous partial nitrification, ANAMMOX and denitrification process in a single reactor for nitrogen removal. 2009, Bioresource Technology; 100: 1548-54.
- Chen J., Weimer P.J. Competition among three predominant cellulolytic bacteria in the absence or presence of non-cellulolytic bacteria. 2001, Microbiology; 147: 21-30 .
- Cheng K.J., Costerton J.W. Ultrastructure of *Butyrivibrio fibrisolvens*: a gram-positive bacterium. 1977, Journal of Bacteriology; 129: 1506-12.
- Choi J.H., Im W.T., Liu Q.M., Yoo J.S., Shin J.H., Rhee S.K., Roh D.H. Planococcus donghaensis sp. nov., a starch-degrading bacterium isolated from the East Sea, South Korea. 2007, International Journal of Systematic and Evolutionary Microbiology; 57: 2645-50.
- Chow C.S., Lamichhane T.N., Mahto S.K. Expanding the nucleotide repertoire of the ribosome with post-transcriptional modifications. 2007, ACS Chemical Biology; 2: 610-9.
- Colwell R.R. Polyphasic taxonomy of the genus vibrio: numerical taxonomy of *Vibrio cholerae*, *Vibrio parahaemolyticus*, and related *Vibrio* species. 1970, Journal of Bacteriology; 104: 410-33.
- Condon C. Maturation and degradation of RNA in bacteria. 2007, Current Opinion in

Microbiology; 10: 271-8.

- Curtis T.P., Sloan W.T., Scannell J.W. Estimating prokaryotic diversity and its limits. 2002, PNAS; 99: 10494-9.
- Czerkawski J. W. (ed.). An introduction to rumen studies. 1988, Pergamon Press, Oxford, England.
- Dalsgaard T, Thamdrup B. Factors controlling anaerobic ammonium oxidation with nitrite in marine sediments. 2002, Applied and Environmental Microbiology; 68: 3802-3808.
- De Liphthay J.R., Ezinger C., Jonsen K., Aamand J., Sorensen S.J. Impact of DNA extraction method on bacterial community composition measured by denaturing gradient gel electrophoresis. 2004, Soil Biology & Biochemistry; 36: 1607-1614,
- De Menezes A.B., Lewis E., O'Donovan M., O'Neill B.F., Clipson N., Doyle E.M.. Microbiome analysis of dairy cows fed pasture or total mixed ration diets. 2011, FEMS Microbiology Ecology; 78: 256-65.
- Delwicke CC. The Nitrogen Cycle. 1970, Scientific American; 223: 1137-46
- Deutscher M.P. Degradation of RNA in bacteria: comparison of mRNA and stable RNA. 2006, Nucleic Acids Research; 34: 659-66.
- Devenish J.A., Barnum D.A.. Evaluation of API 20E System and Encise Enterotube for the identification of Enterobacteriaceae of animal origin. 1980, Canadian Journal of Comparative Medicine and Veterinary Science; 44: 315-9.
- Din N., Damude H.G., Gilkes N.R., Miller R.C., Warren R.A.J., Kilburn D.G. C<sub>1</sub>-C<sub>x</sub> revisited: intramolecular synergism in cellulose. 1994, PNAS; 91: 11383-11387.
- Doerner K.C., White B.A. Assessment of the endo-1,4-beta-glucanase components of *Ruminococcus flavefaciens* FD-1. 1990 Applied and Environmental Microbiology; 56: 1844-50.
- Doi R.H., Kosugi A. Cellulosomes: plant-cell-wall-degrading enzyme complexes. 2004, Nature Reviews Microbiology; 2: 541-51.
- Dunbar J., Takala S., Barns S.M., Davis J.A., Kuske C.R.. Levels of bacterial community diversity in four arid soils compared by cultivation and 16S rRNA gene cloning. 1999, Applied and Environmental Microbiology; 65: 1662-9.
- Duncan S.H., Lobley G.E., Holtrop G., Ince J., Johnstone A.M., Louis P., Flint H.J. Human colonic microbiota associated with diet, obesity and weight loss. 2008, International Journal of Obesity; 32: 1720–1724.
- Egli K, Bosshard F, Werlen C, Lais P, Siegrist H, Zehnder AJB, van der Meer JR. Microbial composition and structure of a rotating biological contactor biofilm treating ammonium-rich wastewater without organic carbon. 2003, Microbial Ecology; 45: 419 - 432 .
- Elston H.R., Baudo J.A., Stanek J.P., Schaab M. Multi-biochemical test system for distinguishing enteric and other gram-negative bacilli. 1971, Applied Microbiology; 22: 408-14.
- Evans C.S., Hedger J.N. Degradation of cell wall polymers. In Gadd G.M. (ed.), Fungi in Bioremediation, 1-26. Cambridge University Press.

- Even S., Pellegrini O., Zig L., Labas V., Vinh J., Bréchemmier-Baey D., Putzer H. Ribonucleases J1 and J2: two novel endoribonucleases in *B. subtilis* with functional homology to E.coli RNase E. 2005, Nucleic Acids Research; 33: 2141-52.
- Evguenieva-Hackenberg E. Bacterial ribosomal RNA in pieces. 2005, Molecular Microbiology; 57: 318-25. Hunt DE, Klepac-Ceraj V, Acinas SG, Gautier C,
- Fantroussi S.E. Verschuere L., Verstraete W. Effect of Phenylurea Herbicides on soil microbial communities estimated by analysis of 16S rRNA gene fingerprintings and community-level physiological profiles. 1999, Applied and Environmental Microbiology; 65: 982-989.
- Fayyaz-ur-Rehman M., Ilyas Tariq M., Aslam M., Khadija G., Iram A. Inhibition studies of cellulolytic activity isolated from *Planococcus citri*. 2009, The Open Enzyme Inhibition Journal; 2:8-11.
- Feinstein L.M., Sul W.J. Blackwood C.B. Assessment of bias associated with incomplete extraction of microbial DNA from soil. 2009, Applied and Environmental Microbiology; 75: 5428-5433.
- Felske A., Engelen B., Nubel U., Backhaus H. Direct ribosome isolation from soil to extract bacterial rRNA for community analysis. 1996, Applied and Environmental Microbiology; 62: 4162-4167.
- Ferris M.J., Muyzer G., Ward D.M., Denaturing gradient gel electrophoresis profiles of 16S rRNA defined populations inhabiting a hot spring microbial mat community. 1996, Applied and Environmental Microbiology; 62: 340-346.
- Firkins J.L., Yu Z., Morrison M. Ruminant nitrogen metabolism: Perspectives for integration of microbiology and nutrition for dairy. 2007, Journal of Dairy Science; 90: E1–E16.
- Flint H.J., Bayer E.A., Rincon M.T., Lamed R., White B.A. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. 2008, Nature Reviews Microbiology; 6: 121-131.
- Francis C.A., Beman J.M., Kuypers M.M.M., New processes and players in the nitrogen cycle: the microbial ecology of anaerobic and archaeal ammonia oxidation. 2007, The ISME Journal; 1: 19-27 .
- Fromin N., Hamelin J., Tarnawski S., Roesti D., Jourdain-Miserez K., Forestier N., Teyssier-Cuvelle S., Gillet F., Aragno M., Rossi P. Statistical analysis of denaturing gel electrophoresis (DGE) fingerprinting patterns. 2002, Environmental Microbiology; 4: 634-643.
- Fujimoto N., Tomoyuki K., Nakao T., Yamada N. *Bacillus licheniformis* Bearing a High Cellulose-Degrading Activity, which was Isolated as a Heat-Resistant and Micro-Aerophilic Microorganism from Bovine Rumen. 2011, The Open Biotechnology Journal; 5: 7-13
- Fuller R. A review: probiotics in man and animals. 1989, Journal of Applied Bacteriology; 66: 803–830.
- Furer G. Wehrli B. Microbial reactions, chemical speciation and multicomponent diffusion in porewaters of a eutrophic lake. 1996, Geochimica et Cosmochimica Acta; 60, 2333-2346 .

- Gao H., Sengupta J., Valle M., Korostelev A., Eswar N., Stagg S.M., Van Roey P., Agrawal R.K., Harvey S.C., Sali A., Chapman M.S., Frank J. Study of the structural dynamics of the E coli 70S ribosome using real-space refinement. 2003, *Cell*; 113: 789-801.
- Gegenheimer P., Watson N., Apirion D. Multiple pathways for primary processing of ribosomal RNA in *Escherichia coli*. 1977, *The Journal of Biological Chemistry* ;252: 3064-73.
- Gevers D., Coan F.M., Lawrence J.G., Spratt B.G., Coenye T., Feil E.J., Stackenbrandt E., Van de Peer Y., Vandamme P., Thompson L., Swings J. Re-Evaluating prokaryotic species. 2005, *Nature Reviews Microbiology*; 3: 733-739.
- Giovannoni S.J., Britschgi T.B., Moyer C.L., Field K.G.. Genetic diversity in Sargasso Sea bacterioplankton. 1990, *Nature*; 45: 60-3.
- Glöckner FO, Kube M, Bauer M, Teeling H, Lombardot T, Ludwig W, Gade D, Beck A, Borzym K, Heitmann K, et al. Complete genome sequence of the marine planctomycete *Pirellula* sp. Strain 1. 2003, *PNAS*; 100: 8298–8303
- Gross K.L., Harmon D.L., Avery T.B. Portal-drained visceral flux of nutrients in lambs fed alfalfa or maintained by total intragastric infusion. 1990, *Journal of Animal Science*; 68: 214-21.
- Gutell R.R., Lee J.C., Cannone J.J. The accuracy of ribosomal RNA comparative structure models. 2002, *Current Opinion in Structural Biology*; 12: 301-10.
- Gutsell N.S., Jain C. Coordinated regulation of 23S rRNA maturation in *Escherichia coli*. 2010, *Journal of Bacteriology*; 192: 1405-9.
- Gürtler V., Stanisich V.A. New approaches to typing and identification of bacteria using the 16S-23S rDNA spacer region. 1996, *Microbiology*; 142: 3-16.
- Güven D, van de Pas-Schoonen K, Schmid MC, Strous M, Jetten MS, Sözen S, Orhon D, Schmidt I. Implementation of the Anammox process for improved nitrogen removal. 2004, *J Environ Sci Health A Tox Hazard Subst Environ Eng.*; 39: 1729-38.
- Güven D, Dapena A, Kartal B, Schmid MC, Maas B, van de Pas-Schoonen K, Sozen S, Mendez R, Op den Camp HJ, Jetten MS, Strous M, Schmidt I. Propionate oxidation by and methanol inhibition of anaerobic ammonium-oxidizing bacteria. 2005, *Applied and Environmental Microbiology*; 71: 1066-71.
- Haddi M.L., Kalid-Meniai K., Kassem M.M., Filacorda S., Susmel P. Comparison between exponential and logistic models for estimation of kinetic parameters of in vitro fermentation of *Atriplex halimus* by mixed rumen microflora of dromedary camel and cattle. 2000, *Annals of Agricultural Science Cairo*; 2: 575-593
- Haddi M.L., Filacorda S., Khalid-Meniai K., Rollin F., Susmel P. In vitro fermentation kinetics of some halophyte shrubs sampled at three stages of maturity. 2003, *Animal Feed Science and Technology*; 104: 215-225.
- Hao X, Heijnen JJ, van Loosdrecht MC. Sensitivity analysis of a biofilm model describing a one-stage completely autotrophic nitrogen removal process. 2002, *Biotechnology and Bioengineering*; 77: 266-77.
- Hashelford K.E., Chuzhanova N.A., Fry J.C., Jones A.J, Weightman A.J. At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain

- substantial anomalies. 2005, *Applied and Environmental Microbiology*; 71: 7724-36.
- Helaszek C.T., White B.A. Cellobiose uptake and metabolism by *Ruminococcus flavefaciens*. 1991, *Applied and Environmental Microbiology*; 57: 64-8.
- Helm M. Post-transcriptional nucleotide modification and alternative folding of RNA. 2006, *Nucleic Acids Research*; 34: 721-33.
- Hendriks J, Oubrie A, Castresana J, Urbani A, Gemeinhardt S, Saraste M. Nitric oxide reductases in bacteria. 2000, *Biochimica et Biophysica Acta*; 1459: 266-73.
- Hentschel U, Hopke J, Horn M, Friedrich AB, Wagner M, Hacker J, Moore BS. Molecular evidence for a uniform microbial community in sponges from different oceans. 2002, *Applied and Environmental Microbiology*; 68: 4431-4440.
- Herschleb J., Ananiev G., Schwartz D.C. Pulsed-field gel electrophoresis. 2007, *Nature Protocols*; 2: 677-84.
- Herskovitz M.A., Bechhofer D.H. Endoribonuclease RNase III is essential in *Bacillus subtilis*. 2000, *Molecular Microbiology*; 38: 1027-33.
- Hespell R.B., Wolf R., Bothast R.J. Fermentation of xylans by *Butyrivibrio fibrisolvens* and other ruminal bacteria. 1987, *Applied and Environmental Microbiology*; 53: 2849-53.
- Hespell R.B. Fermentation of xylans by *Butyrivibrio fibrisolvens* and *Thermoanaerobacter* strain B6A: utilization of uronic acids and xylanolytic activities. 1992, *Current Microbiology*; 25: 189-195.
- Hill T.C., Walsh K.A. Harris J.A., Moffett B.F.. Using ecological diversity measures with bacterial communities. 2003, *FEMS Microbiology Ecology*; 43: 1-11.
- Hino T., Nagano S., Sugimoto H., Tosha T., Shiro J., Molecular structure and function of bacterial nitric oxide reductase. 2011, *Biochimica et Biophysica Acta*; doi:10.1016/j.bbabi.2011.09.021.
- Hopmans EC, Kienhuis MV, Rattray JE, Jaeschke A, Schouten S, Damsté JS. Improved analysis of ladderane lipids in biomass and sediments using high-performance liquid chromatography/atmospheric pressure chemical ionization tandem mass spectrometry. 2006, *Rapid Communications in Mass Spectrometry*; 20: 2099-103.
- Hugenholtz P., Goebel B.M. The polymerase chain reaction as a tool to investigate microbial diversity in environmental samples. 2001, *Environmental Molecular Microbiology: Protocols and Applications*. Rochelle P.A. (ed). Norfolk, UK: Horizon Scientific Press, pp. 31-42.
- Hugenholtz P. Exploring prokaryotic diversity in the genomic era. 2002, *Genome Biology*; 3 reviews0003.1-03.8.
- Hughes J.B., Hellmann J.J., Ricketts T.H., Bohannon B.J. Counting the uncountable: statistical approaches to estimating microbial diversity. 2001, *Applied and Environmental Microbiology*; 67: 4399-406.
- Hume I.D., Warner A.C.I. Evolution of microbial digestion in mammals. 1980, In Y. Ruckebusch and P. Thivend (eds.), *Digestive physiology and metabolism in ruminants*, 665-684. MTD Press, Lancaster, England.
- Hungate R.E.. The anaerobic mesophilic cellulolytic bacteria. 1950, *Bacteriological*

Review; 14: 1-49.

- Hungate R.E. Symposium on “Nutritional Implication of Microbial Action in the non-ruminal alimentary tract”. 1984, Proceeding of the Nutrition Society; 43: 1-11.
- Hunt D.E., Klepac-Ceraj V., Acinas, S.G., Gauthier C., Bertilsson S., Polz M.F. Evaluation of 23S rRNA PCR Primers for use in Phylogenetic Studies of Bacterial Diversity. 2006, Applied Environmental Microbiology; 72: 2221-2225.
- Jaeschke A, Op den Camp HJ, Harhangi H, Klimiuk A, Hopmans EC, Jetten, Schouten S, Sinninghe Damste JS. 16S rRNA gene and lipid biomarker evidence for anaerobic ammonium-oxidizing bacteria in California and Nevada hot springs. 2009, FEMS Microbiology Ecology; 67: 343-350
- Janus HM, van der Roest HF. Don't reject the idea of treating reject water. 1997, Water Science and Technology; 35: 27-34.
- Jetten, M.S.M., Strous M., van de Pas-Schoonen K.T., Schalk J., van Dongen U.G.J.M., van de Graaf, A.A., Logemann S., Muyzer G., van Loosdrecht M.C.M., Kuenen J.G. The anaerobic oxidation of ammonium. 1999, FEMS Microbiology Reviews; 22: 421 - 437 .
- Jetten MS, Wagner M, Fuerst J, van Loosdrecht M, Kuenen G, Strous M. The anammox case-a new experimental manifesto for microbiological eco-physiology. 2001-a, Current Opinion in Biotechnology; 12: 283-8.
- Jetten MSM, Wagner M, Fuerst J, van Loosdrecht M, Kuenen G, Strous M. Microbiology and application of the anaerobic ammonium oxidation (‘anammox’) process. 2001b: Current Opinion in Biotechnology; 12:283 - 288 .
- Jetten MS, Sliemers O, Kuypers M, Dalsgaard T, van Niftrik L, Cirpus I, van de Pas-Schoonen K, Lavik G, Thamdrup B, Le Paslier D, Op den Camp HJ, Hulth S, Nielsen LP, Abma W, Third K, Engström P, Kuenen JG, Jørgensen BB, Canfield DE, Sinninghe Damsté JS, Revsbech NP, Fuerst J, Weissenbach J, Wagner M, Schmidt I, Schmid M, Strous M. Applied Microbiology and Biotechnology. 2003 ;63: 107-14.
- Josaitis C.A., Gaal T., Gourse R.L. Stringent control and growth-rate-dependent control have nonidentical promoter sequence requirements. 1995, PNAS; 92: 1117-21.
- Jouany J.P., Dardillat C., Kayouli C. Microbial cell wall digestion in camelids. 1995, CIHEAM, Options, Seminaire du Projet CEE-DGXII TS2\*0233-C (EDB), Douz, Tunisia.
- Jouany J.P. Effect of rumen protozoa on nitrogen utilization by ruminants. 1996, Journal of Nutrition; 126: 1335-46.
- Jun H.S., Qi M., Ha J.K., Forsberg C.W. *Fibrobacter succinogenes*, a dominant fibrolytic ruminal bacterium: transition to the post genomic era. 2007a, The Asian-Australasian Association of Animal Production Societies; 20: 802-810.
- Jun H.S., Qi M., Gong J., Egbosimba E.E., Forsberg C.W.. Outer membrane proteins of *Fibrobacter succinogenes* with potential roles in adhesion to cellulose and in cellulose digestion. 2007b, Journal of Bacteriology; 189: 6806-15.
- Jun S.R., Sims G.E., Wu G.A., Kim S.H. Whole-proteome phylogeny of prokaryotes by feature frequency profiles: An alignment-free method with optimal feature resolution. 2010, PNAS; 107: 133-8.
- Kartal B, Rattray J, van Niftrik LA, van de Vossenberg J, Schmid MC, Webb RI,

- Schouten S, Fuerst JA, Damsté JS, Jetten MS, Strous M. Candidatus "*Anammoxoglobus propionicus*" a new propionate oxidizing species of anaerobic ammonium oxidizing bacteria. 2007, *Systematic and Applied Microbiology*; 30: 39-49.
- Kayouli C., Jouani J.P., Demeyer D.I., Ali-Ali, Taoueb H., Dardillat C, 1993. Corporative studies on the degradation and the mean retention time of solid and liquid phases in the foregut of dromedaries and sheep fed on low-quality roughages from Tunisia. 1993, *Animal Feed Science and Technology*; 40: 343-355.
  - Kelly W.J., Leahy S.C., Altermann E., Yeoman C.J., Dunne J.C., Kong Z., Pacheco D.M, Li D., Noel S.J, Moon C.D., Cookson A.L., Attwood G.T. The glycobiome of the rumen bacterium *Butyrivibrio proteoclasticus* B316 highlights adaptation to a polysaccharide-rich environment. 2010, *PLoS One*; 5: e11942.
  - Kenneth L., Heck J., van Belle G., Simberloff D. Explicit Calculation of the Rarefaction Diversity Measurement and the Determination of Sufficient sample size. 1975, *Ecology*; 56: 1459-1461.
  - Kim M., Morrison M., Yu Z. Status of the phylogenetic diversity census of ruminal microbiomes. 2010 *FEMS Microbiology Ecology*; 76: 49-63.
  - Klappenbach J.A., Dunbar J.M., Schmidt T.M. rRNA operon copy number reflects ecological strategies of bacteria. 2000, *Applied and Environmental Microbiology*; 66: 1328-33.
  - Kobayashi Y., Shinkai T., Koike S.. Ecological and physiological characterization shows that *Fibrobacter succinogenes* is important in rumen fiber digestion-review. 2008, *Folia Microbiologica* ; 53: 195-200.
  - Koellner T., Hersperger A. M., Wohlgemuth T. Rarefaction method for assessing plant species diversity on a regional scale. 2004, *Ecography*; 27: 532–544.
  - Koenig K.M., Newbold C.J., McIntosh F.M., Rode L.M. Effects of protozoa on bacterial nitrogen recycling in the rumen. 2000, *Journal of Animal Science*; 78: 2431-45.
  - Koike S. Kobayashi Y. Fibrolitic rumen bacteria: their ecology and functions. 2009, *The Asian-Australasian Association of Animal Production Societies*; 22: 131-138.
  - Koonin E.V., Deutscher M.P. RNase T shares conserved sequence motifs with DNA proofreading exonucleases. 1993, *Nucleic Acids Research*; 21: 2521-2.
  - Kopke B., Wilms R., Englen B., Cypionka H., Sass H.. Microbial diversity in coastal subsurface sediments: a cultivation approach using various electron acceptors and substrate gradients. 2004, *Applied Environmental Microbiology*; 71: 7819-7830.
  - Krehbiel, C. R.; Rust, S. R.; Zhang, G., 2003: Bacterial direct-fed microbials in ruminant diets: performance response and mode of action. *Journal of Animal Science* 81(E. Suppl. 2), 120–132.
  - Kristensen N.B., Gabel G., Pierzynowski S.G., Danfaer A. Portal recovery of short-chain fatty acids infused into the temporarily-isolated and washed reticulorumen of sheep. 2000 *British Journal of Nutrition*, 84: 477-482.
  - Kristensen N.B., Harmon D.L. Splanchnic metabolism of volatile fatty acids absorbed from the washed reticulorumen. 2004, *Journal of Animal Science*, 82: 2033-2042.
  - Kuai L, Verstraete W. Ammonium removal by the oxygen-limited autotrophic nitrification-denitrification system. 1998, *Applied and Environmental Microbiology*; 64:

4500-6.

- Kuenen, J.G., Anammox bacteria: from discovery to application. 1998, *Nature Reviews Microbiology*; 6: 320 - 326 .
- Kumar M, Lin JG. Co-existence of anammox and denitrification for simultaneous nitrogen and carbon removal--Strategies and issues. 2010, *Journal of Hazardous Materials*; 178: 1-9.
- Kunin V., Goldovsky L., Darzentas N., Ouzounis C.A.. The net of life: reconstructing the microbial phylogenetic network. 2005, *Genome Research*; 15: 954-9.
- Kuypers, M.M.M., Sliemers, A.O., Lavik, G., Schmid, M., Jørgensen, B.B., Kuenen, J.G., Sinninghe Damstè, J.S., Strous, M., Jetten, M.S.M., Anaerobic ammonium oxidation by anammox bacteria in the Black Sea. 2003, *Nature*; 422 608-611 .
- Lafontaine DL, Tollervey D. The function and synthesis of ribosomes. *Nature Reviews Molecular Cell Biology* 2001 ;2: 514-20.
- Lamed R., Setter E., Bayer E.A. Characterization of a cellulose-binding, cellulase-containing complex in *Clostridium thermocellum*. 1983, *Journal of Bacteriology*; 156: 828-36.
- Lane D.J, Pace B., Olsen G.J., Stahl D.A., Sogin M.L., Pace N.R. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. 1985, *PNAS*; 82: 6955-9.
- Larue R., Yu Z., Parisi V.A., Egan A.R., Morrison M. Novel microbial diversity adherent to plant biomass in the herbivore gastrointestinal tract, as revealed by ribosomal intergenic spacer analysis and rrs gene sequencing. 2005, *Environmental Microbiology*; 7: 530-43.
- Lechner-Doll M., Kaske M., Engelhardt W.V. Factors affecting the mean retention time of particles in the forestomach of ruminant and camelids. In: *Physiological Aspects of Digestion and Metabolism in Ruminants*, 1991. Tsuda T., Sasaki Y., Kawashima R., (eds), Academic Press New-York, pp 455-482.
- Lee C.Y., Lee J.C., Gutell R.R. Networks of interactions in the secondary and tertiary structure of ribosomal RNA. 2007, *Physica Acta*; 386: 564-572.
- Lee Z.M., Bussema C., Schmidt T.M. rrnDB: documenting the number of rRNA and tRNA genes in bacteria and archaea. 2008, *Nucleic Acids Research*; 37: D489-93.
- Leng, R.A., Nolan J.V. Nitrogen metabolism in the rumen. 1984, *Journal of Dairy Science* 67: 1072–1089.
- Leng J., Xie L., Zhu R., Yang S., Gou X., Li S., Mao H. Dominant bacterial communities in the rumen of Gayals (*Bos frontalis*), Yaks (*Bos grunniens*) and Yunnan Yellow Cattle (*Bos taurus*) revealed by denaturing gradient gel electrophoresis. 2011, *Molecular Biology Reports*; 38: 4863-72.
- Leontis N.B., Westhof E. Geometric nomenclature and classification of RNA base pairs. 2001, *RNA*; 7: 499-512.
- Li R.W., Connor E.E., Li C., Baldwin R.L., Sparks M.E. Characterization of the rumen microbiota of pre-ruminant calves using metagenomic tools. 2011, *Environmental Microbiology*; 14: 129-39.
- Li Z., Deutscher M.P. The tRNA processing enzyme RNase T is essential for maturation

of 5S RNA. 1995, PNAS; 92: 6883-6.

- Li Z., Pandit S., Deutscher M.P. Maturation of 23S ribosomal RNA requires the exoribonuclease RNase T. 1999, RNA; 5: 139-46.
- Liiv A., O'Connor M. Mutations in the intersubunit bridge regions of 23 S rRNA. 2006, The Journal of Biological Chemistry; 281: 29850-62.
- Lindsay MR, Webb RI, Fuerst JA. Pirellulosomes: a new type of membrane-bounded cell compartment in planctomycete bacteria of the genus Pirellula. 1997, Microbiology; 143: 739-48
- Lindsay MR, Webb RI, Strous M, Jetten MS, Butler MK, Forde RJ, Fuerst JA. Cell compartmentalisation in planctomycetes: novel types of structural organisation for the bacterial cell. 2001, Archives of Microbiology; 175: 413-29.
- Liu Z., Lozupone C., Hamady M., Bushman F.D., Knight R. Short pyrosequencing reads suffice for accurate microbial community analysis. 2007, Nucleic Acids Research; 35: e120.
- Lodish H., Berk A., Zipursky S. L., Matsudaira P., Baltimore D., Darnell J. Molecular Cell Biology, 4th edition. 2000 Par. 11.5. New York: W. H. Freeman. (Source: ncbi.nlm.nih.gov/books/NBK21475)
- Loste A., Ramos J.J., Garcia L., Ferrer L.M., Verde M.T. High prevalence of ulcerative in Rasa Aragonesa Rams Associated with a legume-rich diet. 2005, The Journal of Veterinary Medical Science; 52: 176-179.
- Lynd L.R., Weimer P.J., van Zyl W.H., Pretorius I.S., Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS. Microbial cellulose utilization: fundamentals and biotechnology. Microbiology and Molecular Biology Reviews 2002 ;66: 506-77, table of contents.
- Mackie R.I. Mutualistic fermentative digestion in the gastrointestinal tract: diversity and evolution. 2002, Integr Comp Biol.; 42: 319-26.
- Mahowald MA, Rey FE, Seedorf H, Turnbaugh PJ, Fulton RS, Wollam A, Shah N, Wang C, Magrini V, Wilson RK, Cantarel BL, Coutinho PM, Henrissat B, Crock LW, Russell A, Verberkmoes NC, Hettich RL, Gordon JI. Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. 2009 PNAS; 106: 5859-64.
- Mares M.A. Encyclopedia of Deserts. 1999, University of Oklahoma Press, p 183.
- Mathy N., Bénard L., Pellegrini O., Daou R., Wen T., Condon C. 5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of Mrna. 2007, Cell; 129: 681-92.
- Meroueh M., Grohar P.J., Qiu J., SantaLucia J.Jr., Scaringe S.A., Chow C.S. Unique structural and stabilizing roles for the individual pseudouridine residues in the 1920 region of *Escherichia coli* 23S rRNA. 2000, Nucleic Acids Research; 28: 2075-83.
- McClory S.P., Leisring J.M., Qin D, Fredrik K. Missense suppressor mutation in 16S rRNA reveal the importance of helices h8 and h14 in aminoacyl-tRNA selection. 2010, RNA; 16: 1925-1934.
- Michalet-Doreau B., Fernandez I., Peyron C., Millet L., Fonty G. Fibrolytic activities and cellulolytic bacterial community structure in the solid and liquid phases of rumen contents. 2001, Reproduction Nutrition Development; 41: 187-94.

- Miller T.L. The pathway of formation of acetate and succinate from pyruvate by *Bacteroides succinogenes*. 1978, Archives of Microbiology; 117: 145-152.
- Miller M.B., Antonopoulos D.A., Rincon M.T., Band N., Bari A., Akraiko T., Hernandez A., Thimmapuram J., Henrissat B., Coutinho P.M, Borovok I., Jindou S., Lamed R., Flint H.J., Byer E.A., White B.A. Diversity and strain specificity of plant cell wall degrading enzymes revealed by the draft genomes of *Ruminococcus Flavefaciens* FD-1. 2009, Plos ONE, 4: e6650.
- Miron J., Ben-Ghedalia D., Morrison M. Invited review: adhesion mechanisms of rumen cellulolytic bacteria. 2001 Journal of Dairy Science; 84: 1294-309.
- Moat A.G., Foster J.W., Spector M.P. Microbial Physiology, 4<sup>th</sup> Edition, 2002. John Wiley & Sons, Inc , publication, pp. 488-489.
- Mohammed R., Stevenson D.M., Beauchemin K.A., Muck R.E., Weimer P.J. Changes in ruminal bacterial community composition following feeding of alfalfa ensiled with a lactic acid bacterial inoculant. 2012, Journal of Dairy Science; 95: 328-339.
- Monnet C., Loux V., Bento P., Gibrat J.F., Straub C., Bonnarme P., Landaud S., Irlinger F. Genome Sequence of *Corynebacterium casei* UCMA 3821, Isolated from a Smear-Ripened Cheese. 2012, Journal of Bacteriology; 194: 738-9.
- Mounier J., Rea M.C., O'Connor P.M., Fitzgerald G.F., Cogan T.M. Growth characteristics of *Brevibacterium*, *Corynebacterium*, *Microbacterium*, and *Staphylococcus spp.* isolated from surface-ripened cheese. 2007, Applied and Environmental Microbiology; 73: 7732-9.
- Moura I, Moura J.J. Structural aspects of denitrifying enzymes. 2001 Current Opinion in Chemical Biology; 5:168-75.
- Mulder A., van de Graaf A.A., Robertson L.A., Kuenen J.K Anaerobic ammonium oxidation discovered in a denitrifying fluidized bed reactor. 1995 FEMS Microbiology Ecology; 16: 177-184.
- Mulder A., The quest for sustainable nitrogen removal technologies. 2003, Water Science and Technology; 48: 67-75.
- Murray H.D., Appleman J.A., Gourse R.L. Regulation of the *Escherichia coli* *rrnB* P2 promoter. 2003, Journal of Bacteriology; 185:28-34.
- Muyzer G., de Waal E.C., Uitterlinden A.G. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S Rrna. 1993 Applied and Environmental Microbiology; 59: 695-700.
- Myers R.M., Fischer S.G., Lerman L.S., Maniatis T. Nearly all single base substitutions in DNA fragments joined to a GC-clamp can be detected by denaturing gradient gel electrophoresis. 1985 Nucleic Acids Research; 13: 3131-45.
- Navarre, W. W., and O. Schneewind. Proteolytic cleavage and cell wall anchoring at the LPXTG motif of surface proteins in gram-positive bacteria. Molecular Microbiology 14: 115-121, 1994.
- Neef A., Arnann R., Schlesner H., Schleifer H.K. Monitoring a widespread bacterial

- group: in situ detection of Planctomycetes with 16S rRNA-targeted probes. 1998, *Microbiology*; 144: 3257-3266.
- Nierhaus K.H, Wilson D.N. (Eds.). Protein synthesis and ribosome structure. Translating the genome. 2004, Wiley-VCH; 429-430.
  - Nojiri M., Xie Y., Inoue T., Yamamoto T, Matsumura H., Kataoka K., Deligeer, Yamaguchi K., Kai Y., Suzuki S. Structure and function of a hexameric copper-containing nitrite reductase. 2007, *PNAS*; 104: 4315-20.
  - Nolan J.V., Dobos R.C. Nitrogen transactions in ruminants. 2005, Pages 177–206 in *Quantitative Aspects of Ruminant Digestion and Metabolism*. 2nd. ed. J. Dijkstra, J. M. Forbes, and J. France, ed. CABI Publ., Wallingford, UK.
  - Noller H.F. Biochemical characterization of the ribosomal decoding site. 2006, *Biochimie*; 88: 935-41.
  - Norton, J. M., J. J. Alzerreca, Y. Suwa, and M. G. Klotz. Diversity of ammonia monooxygenase operon in autotrophic ammonia-oxidizing bacteria. 2002, *Archives of Microbiology*; 177: 139–149 .
  - Nubel U., Engelen B., Felske A., Snaird J., Wieshuber A., Amman R., Ludwig W, Backhaus H. Sequence heterogeneities of genes encoding 16S rRNAs in *Paenibacillus polymixa* detected by temperature gradient gel electrophoresis. 1996, *Journal of Bacteriology*, 178: 5636-5643.
  - Numata, M., Saito T., Yamazaki, Y. Dukumori Y., Yamanaka T. Cytochrome P-460 of *Nitrosomonas europaea*: further purification and further characterization. *The Journal of Biochemistry*. 108: 1016–1023 .
  - Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. *Nature*. 2000 ;405: 299-304.
  - Ogle J.M., Ramakrishnan V. Structural insights into translational fidelity. 2005, *Annual Review of Biochemistry*; 74: 129-77.
  - Okabe S., Oshiki M., Kamagata Y., Yamaguchi N., Toyofuku M., Yawata Y., Tashiro Y., Nomura N., Ohta H., Ohkuma M., Hiraiishi A., Minamisawa K. A great leap forward in microbial ecology. 2010, *Microbes and Environments*; 25: 230-40.
  - Olmedo G., Guzmán P. Mini-III, a fourth class of RNase III catalyses maturation of the *Bacillus subtilis* 23S ribosomal RNA. 2008, *Molecular Microbiology*; 68: 1073-6.
  - Op den Camp H.J., Kartal B., Guven D., van Niftrik L.A., Haaijer S.C., van der Star W.R., van de Pas-Schoonen K.T., Cabezas A., Ying Z., Schmid M.C., Kuypers M.M., van de Vossenberg J., Harhangi H.R., Picioreanu C., van Loosdrecht M.C., Kuenen J.G., Strous M., Jetten M.S. Global impact and application of the anaerobic ammonium-oxidizing bacteria. 2006, *Biochemical Society Transactions*; 34: 174-8.
  - Otsuka J., Terai G., Nakano T. Phylogeny of organisms investigated by the base-pair changes in the stem regions of small and large ribosomal subunit RNAs. 1999, *Journal of Molecular Evolution*; 48: 218-235.
  - Ow M.C., Perwez T., Kushner S.R. RNase G of *Escherichia coli* exhibits only limited functional overlap with its essential homologue RNase E. 2003, *Molecular Microbiology*; 49: 607-22.
  - Pace NR. A molecular view of Microbial Diversity and the Biosphere. 1997, *Science*;

276: 734-740.

- Park I, Jaiesoon C. Partial characterization of extracellular xylanolytic activity derived from *Paenibacillus* sp. KIJ1. 2010, African Journal of Microbiology Research; 4: 1257-1264.
- Parker D.S., Lomax M.A., Seal C.J., Wilton J.C. Metabolic implications of ammonia production in the ruminant. 1995, Proceedings of the Nutrition Society; 54: 549-563.
- Pason P., Kyu K.L., Ratanakhanokchai K. *Paenibacillus curdlanolyticus* strain B-6 xylanolytic-cellulolytic enzyme system that degrades insoluble polysaccharides. 2006, Applied and Environmental Microbiology; 72: 2483-90.
- Pauly M., Keegstra K. Cell-wall carbohydrates and their modification as a resource for biofuels. 2008, The Plant Journal; 54: 559-68.
- Pope PB, Denman SE, Jones M, Tringe SG, Barry K, Malfatti SA, McHardy AC, Cheng JF, Hugenholtz P, McSweeney CS, Morrison M. Adaptation to herbivory by the Tammar wallaby includes bacterial and glycoside hydrolase profiles different from other herbivores. PNAS 2010 ;107: 14793-8.
- Pulk A., Maivali U., Remme J. Identification of nucleotides in *E. coli* 16S rRNA essential for ribosome subunit association. 2006, RNA; 12: 790-796.
- Purkhold U., Pommerening-Röser A., Juretschko S., Schmid M.C., Koops H.P., Wagner M. Phylogeny of all recognized species of ammonia oxidizers based on comparative 16S rRNA and amoA sequence analysis: implications for molecular diversity surveys. 2000, Applied and Environmental Microbiology; 66: 5368-82.
- Pynaert K, Smets BF, Beheydt D, Verstraete W. Start-up of autotrophic nitrogen removal reactors via sequential biocatalyst addition. Environmental Science & Technology 2004 Feb 15;38: 1228-35.
- Qi M., Nelson K.E., Daugherty S.C., Nelson W.C., Hance I.R., Morrison M., Forsberg C.W. Genomic differences between *Fibrobacter succinogenes* S85 and *Fibrobacter intestinalis* DR7, identified by suppression subtractive hybridization. Applied and Environmental Microbiology 2008 Feb;74: 987-93.
- Qiao G.H., Shan A.S., Ma N., Ma Q.Q., Sun Z.W. Effect of supplemental *Bacillus* cultures on rumen fermentation and milk yield in Chinese Holstein cows. 2010, Journal of Animal Physiology and Animal Nutrition; 94: 429-436.
- Quan, Z.\_X., Rhee, S.-K., Zuo, J.-E, Yang, Y., Bae, J.-W. Park, J.R., Lee, S.-T., Park, Y.-H., Diversity of ammonium-oxidizing bacteria in a granular sludge anaerobic ammonium-oxidizing reactor. 2008, Environmental Microbiology; 10: 3130-3139.
- Qiu YL, Sekiguchi Y, Imachi H, Kamagata Y, Tseng IC, Cheng SS, Ohashi A, Harada H. *Sporotomaculum syntrophicum* sp. Nov., a novel anaerobic, syntrophic, benzoate-degrading bacterium isolated from methanogenic sludge treating wastewater from teraphthalate manufacturing. 2003, Archives of Microbiology; 179: 242-249
- Wallace R.J., Onodera R., Cotta M.A. Metabolism of nitrogen-containing compounds, 1997, Hobson P.N., Stewart C.S.(Eds.), The Rumen Microbial Ecosystem (2nd ed.), Chapman & Hall, London, UK, pp. 283-328.
- Ramos S., Tejido M.L., Martínez M.E., Ranilla M.J., Carro M.D. Microbial protein synthesis, ruminal digestion, microbial populations, and nitrogen balance in sheep fed

- diets varying in forage-to-concentrate ratio and type of forage. 2009, *Journal of Animal Science*; 87: 2924-2934.
- Rappé M.S, Giovannoni S.J. The uncultured Microbial Majority. 2003, *Annual Review of Microbiology*; 57: 369-94.
  - Redko Y., Bechhofer D.H., Condon C. Mini-III, an unusual member of the RNase III family of enzymes, catalyses 23S ribosomal RNA maturation in *B. subtilis*. 2008, *Molecular Microbiology*; 68: 1096-106.
  - Reynolds, C. K. Quantitative aspects of liver metabolism in ruminants. Pages 351–371 in *Ruminant Physiology: Digestion, Metabolism, Growth and Reproduction*. 1995, Proc. 8th Int. Symp. Rumin. Physiol. W. van Engelhardt, S. Leonhard-Marek, G. Breves, and D. Giesecke, ed. Ferdinand Enke Verlag, Stuttgart, Germany.
  - Reynolds C.K. Economics of visceral energy metabolism in ruminants: Toll keeping or internal revenue service?. 2002, *American Society of Animal Science*; 80: 74-84.
  - Reynolds C.K., Kristensen N.B. Nitrogen recycling through the gut and the nitrogen economy of ruminants: an asynchronous symbiosis. 2008, *American Society of Animal Science*; 86: 293-305.
  - Richardson DJ, Berks BC, Russell DA, Spiro S, Taylor CJ. Functional, biochemical and genetic diversity of prokaryotic nitrate reductases. 2001, *Cellular and Molecular Life Sciences*; 58: 165-78.
  - Rincon M.T., Cepeljnik T., Martin J.C., Lamed R., Barak Y., Bayer E.A., Flint H.J. Unconventional mode of attachment of the *Ruminococcus flavefaciens* cellulosome to the cell surface. 2005, *Journal of Bacteriology*; 187: 7569-78.
  - Rodnina M.V., Daviter T., Gromadski K., Wintermeyer W. Structural dynamics of ribosomal RNA during decoding on the ribosome. 2002, *Biochimie*; 84: 745-54.
  - Rostron W.M., Stuckey D.C., Young A.A. Nitrification of high strength ammonia wastewater: comparative study of immobilization media, 2001, *Water Research*; 35, 1169-1178 .
  - Roy M.K., Singh B., Ray B.K., Apirion D. Maturation of 5-S rRNA: ribonuclease E cleavages and their dependence on precursor sequences. 1983, *The Journal of Biochemistry*; 131: 119-27.
  - Russell J.B., O'Connor J.D., Fox D.G., Van Soest P.J., Sniffen C.J. A net carbohydrate and protein system for evaluating cattle diets: I. Ruminant fermentation. 1992, *Journal of Animal Science*; 70: 3551-61.
  - Russell JB, Rychlik JL. Factors that alter rumen microbial ecology. *Science*. 2001;292: 1119-22.
  - Rysgaard, S., and R. N. Glud. 2004. Anaerobic N<sub>2</sub> production in Arctic sea ice. *Limnology and Oceanography* 49: 86-94.
  - Rzhetsky A. Estimating substitution rates in ribosomal RNA genes. 1995, *Genetics*; 141: 771-783.
  - Sacchi CT, Whitney AM, Mayer LW, Morey R, Steigerwalt A, Boras A, Weyant RS, Popovic T. Sequencing of 16S rRNA gene: a rapid tool for identification of *Bacillus*

anthracis. *Emerging Infectious Diseases* journal 2002 ;8: 1117-23.

- Salmaso R., Valmasoni G. *Energia da biomasse, tra presente e futuro*. 2006, Papergraf Ed. Padova.
- Saluzzi L., Smith A., Stewart C.S. Analysis of bacterial phospholipid markers and plant monosaccharides during forage degradation by *Ruminococcus flavefaciens* and *Fibrobacter succinogenes* in co-culture. 1993, *Journal of General Microbiology*; 139: 2865-2873.
- Samsudin A.A., Evans P.N., Wright A.D., Al Jassim R. Molecular diversity of the foregut bacteria community in the dromedary camel (*Camelus dromedarius*). 2011, *Environmental Microbiology*; 13: 3024-35
- Schleifer KH. Classification of Bacteria and Archaea: past, present and future. *Systematic and Applied Microbiology*. 2009 ;32: 533-42.
- Schloss P.D., Westcott S.L. Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. 2011 *Applied and Environmental Microbiology*; 77: 3219-26.
- Schmid M, Twachtmann U, Klein M, Strous M, Juretschko S, Jetten M, Metzger JW, Schleifer KH, Wagner M. Molecular evidence for genus level diversity of bacteria capable of catalyzing anaerobic ammonium oxidation. 2000, *Systematic and Applied Microbiology*; 23: 93-106.
- Schmid, M., Walsh, K., Webb, R.I., Rijpstra, W.I., van de Pas-Schoonen, K., Verbruggen, M.J., Hill, T., Moffett, B., Fuerst, J.A., Schouten, S., Sinninghe Damsté, J., Harris, J., Shaw, P., Jetten, M.S.M., Strous, M., Candidatus “*Scalindua brodae*”, sp.nov., Candidatus “*Scalindua wagneri*”, sp. nov., two new species of anaerobic ammonium oxidizing bacteria. 2003, *Systematic and Applied Microbiology*; 26: 529 - 538 .
- Schneewind O., Fowler A., Faull K.F.. Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*. 1995, *Science*; 268: 103-6.
- Schneider D.A., Ross W., Gourse R.L. Control of rRNA expression in *Escherichia coli*. 2003, *Current Opinion in Microbiology*; 6: 151-6.
- Schwartz D.C., Cantor C.R. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. 1984, *Cell*; 37: 67-75.
- Schwieger F., Tebbe C. A new approach to utilize PCR-Single-Strand-Conformation Polymorphism for 16S rRNA gene-based microbial community analysis. 1998, *Applied Environmental Microbiology*; 64: 4870-4876.
- Shalck, J., de Vries, S., Kuenen, J.G., Jetten, M.S.M., Involvement of a Novel Hydroxylamine Oxidoreductase in Anaerobic Ammonium Oxidation. 2000, *Biochemistry*; 39: 5405 - 5412 .
- Shinkai T., Kobaiashi Y. Localization of ruminal cellulolytic bacteria on plant fibrous materials as determined by fluorescence in situ hybridization and real-time PCR. 2007, *Applied and Environmental Microbiology*; 73: 1646-1652.
- Shinzato N., Muramatsu M., Matsui T., Watanabe Y. Molecular phylogenetic diversity of the bacterial community in the gut of the termite *Coptotermes formosanus*. 2005, *Bioscience Biotechnology Biochemistry*; 69: 1145–1155
- Shrestha P.M., Noll M., Liesack W. Phylogenetic identity, growth-response time and

- rRNA operon copy number of soil bacteria indicate different stages of community succession. 2007, *Environmental Microbiology*; 9: 2464-74.
- Schuwirth B. Structure of the bacterial ribosome at 3.5 Å Resolution. 2005, *Science*; 310: 827-834.
  - Sijpesteijn A.K. On *Ruminococcus flavefaciens*, a cellulose-decomposing bacterium from the rumen of sheep and cattle. 1951, *Journal of General Microbiology*; 5: 869, 879.
  - Sinninghe Damsté JS, Strous M, Rijpstra WI, Hopmans EC, Geenevasen JA, van Duin AC, van Niftrik LA, Jetten MS. Linearly concatenated cyclobutane lipids form a dense bacterial membrane. *Nature*. 2002 ;419: 708-12.
  - Sinninghe Damsté JS, Rijpstra WI, Geenevasen JA, Strous M, Jetten MS. Structural identification of ladderane and other membrane lipids of planctomycetes capable of anaerobic ammonium oxidation . *FEBS J*. 2005 ;272(16): 4270-83.
  - Sliemers AO, Derwort N, Gomez JL, Strous M, Kuenen JG, Jetten MS. Completely autotrophic nitrogen removal over nitrite in one single reactor. Sliemers AO, Derwort N, Gomez JL, Strous M, Kuenen JG, Jetten MS. Completely autotrophic nitrogen removal over nitrite in one single reactor. 2002, *Water Research*; 36: 2475-2482.
  - Smit E., Leefang P., Wernars K. Detection of shifts in microbial community structure and diversity in soil caused by copper contamination using amplified ribosomal DNA restriction analysis 1997, *FEMS Microbiology Ecology*; 23: 249–261
  - Smit S., Widmann J., Knight R. Evolutionary rates vary among rRNA structural elements. 2007, *Nucleic Acids Research*; 35: 3339-54.
  - Smith T.F., Lee J.C., Gutell R.R., Hartman H. The origin and evolution of the ribosome. 2008, *Biology Direct*; 3: 16.
  - Sogin S.J., Sogin M.L., Woese C.R. Phylogenetic measurement in prokaryotes by primary structural characterization. 1971, *Journal of Molecular Evolution*; 1: 173-84.
  - Sogin M.L., Pace N.R.. In vitro maturation of precursors of 5S ribosomal RNA from *Bacillus subtilis*. 1974, *Nature*; 252: 598-600.
  - Song W.S., Lee M., Lee K. RNase G participates in processing of the 5'-end of 23S ribosomal RNA. 2011, *The Journal of Microbiology*;49: 508-11.
  - Sorek R., Zhu Y., Creevey C.J., Francino M.P., Bork P., Rubin E.M.Genome-wide experimental determination of barriers to horizontal gene transfer. 2007, *Science*; 318: 1449-52.
  - Spieck E., Bock E. The Lithoautotrophic Nitrite-Oxidizing Bacteria. *Bergey's Manual® of Systematic Bacteriology*. 149-153,
  - Squartini A. Where the bugs are: analyzing distribution of bacterial phyla by descriptor keyword search in the nucleotide database, *Microbial Informatics and Experimentation*, Vol 1, No. 7 .
  - Stahl D.A., Pacet B., Marsh T., Pace N.R. The ribonucleoprotein substrate for a ribosomal RNA-processing nuclease. 1984, *The Journal of Biological Chemistry*; 18: 11448-53.

- Staley J.T. Konopka A: measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. 1985, Annual Review of Microbiology; 39: 321-346
- Starkenburg S.R., Larimer F.W, Stein L.Y., Klotz M.G, Chain P.S, Sayavedra-Soto L.A., Poret-Peterson A.T., Gentry M.E, Arp DJ, Ward B, Bottomley PJ. Complete genome sequence of *Nitrobacter hamburgensis* X14 and comparative genomic analysis of species within the genus *Nitrobacter*. 2008, Applied and Environmental Microbiology; 74: 2852-63.
- Starkenburg S.R., Chain P.S., Sayavedra-Soto L.A., Hauser L., Land M.L., Larimer F.W., Malfatti S.A., Klotz M.G., Bottomley P.J., Arp D.J., Hickey W.J. Genome sequence of the chemolithoautotrophic nitrite-oxidizing bacterium *Nitrobacter winogradskyi* Nb-255. 2006, Applied and Environmental Microbiology; 72: 2050-63.
- Stein JL, Marsh TL, Wu KY, Shizuya H, DeLong EF. Characterization of uncultivated prokaryotes: isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon. Journal of Bacteriology 1996 Feb;178: 591-9.
- Stevenson B.S., Schmidt T.M. Life history implications of rRNA gene copy number in *Escherichia coli*. 2004, Applied and Environmental Microbiology; 70: 6670-7.
- Stoodley P., Sauer K., Davies D.G., Costerton JW. Biofilms as complex differentiated communities. 2002, Annual Review of Microbiology; 56: 187-209.
- M, Pelletier E, Mangenot S, Rattei T, Lehner A, Taylor MW, Horn M, et al. Deciphering the evolution and metabolism of an anammox bacterium from a community genome. 2006, Nature; 440: 790-4.
- Strous M.J., Kuenen G., Jetten M.S.M. 1999. Key physiology of anaerobic ammonium oxidation. Applied and Environmental Microbiology 65: 3248-3250.
- Suen G, Weimer PJ, Stevenson DM, Aylward FO, Boyum J, Deneke J, Drinkwater C, Ivanova NN, Mikhailova N, Chertkov O, Goodwin LA, Currie CR, Mead D, Brumm PJ. The complete genome sequence of *Fibrobacter succinogenes* S85 reveals a cellulolytic and metabolic specialist. 2011, PLoS One; 6: e18814.
- Sun Q., Vila-Sanjurjo A., O'Connor M. Mutations in the intersubunit bridge regions of 16S rRNA affect decoding and subunit-subunit interactions on the 70S ribosome. 2010, Nucleic Acids Research; 39: 3321-30.
- Sutton J.D. Digestion and Absorption of Energy Substrates in the Lactating Cow. 1985, Journal of Dairy Science, Volume 68, Issue 12: 3376–3393.
- Sutton M.A., Dragosistis U., Tang Y.S., Fowler D. Ammonia emissions from non-agricultural sources in the UK. 2000, Atmospheric Environment; 34, 855-869 .
- Szymanski M., Barciszewska M.Z., Erdmann V.A., Barciszewski J. 5S rRNA: structure and interactions. 2003, Biochemical Journal; 371: 641–651.
- T. A. McAllister, H. D. Bae, G. A. Jones and K. J. Cheng. Microbial attachment and feed digestion in the rumen. Journal of Animal Science, Vol 72, Issue 11 3004-3018
- Tajima K., Aminov R. I., Nagamine T., Ogata K., Nakamura M., Matsui H., Benno Y. Rumen bacterial diversity as determined by sequence analysis of 16S rDNA libraries. 1999, FEMS Microbiology Ecology; 29: 159–169.

- Tajima K., Aminov R.I., Nagamine T., Matsui H., Nakamura M., Benno Y. Diet dependent shifts in the bacterial population of the rumen revealed with real-time PCR. 2001, *Applied and Environmental Microbiology*; 67: 2766-2774.
- Tamminga S., Protein degradation in the forestomach of ruminants. 1979, *Journal of Animal Science*; 49: 1615-1630
- Tenover F.C., Arbeit R.D., Goering R.V., Mickelsen P.A., Murray B.E., Persing D.H., Swaminathan B. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. 1995, *Journal of Clinical Microbiology*; 33: 2233-9.
- Teske A., Alm E., Regan J.M., Toze S., Rittmann B.E., Stahl D.A. Evolutionary relationships among ammonia- and nitrite-oxidizing bacteria. 1994, *Journal of Bacteriology*; 176: 6623-6630.
- Thakuria D., Schmidt O., Siurain M.M., Egan D., Doohan F.M. Importance of DNA quality in comparative soil microbial community structure analyses. 2008, *Soil Biology and Biochemistry*; 40: 1390-1403.
- Martinez-Murcia A.J., Acinas S.G., Rodriguez-Valera F. Evaluation of prokaryotic diversity by restrictionase digestion of 16S rDNA directly amplified from hypersaline environments. 1995, *FEMS Microbiology Ecology*; 17: 247-255
- Torsvik V., Salte K., Sørheim R., Goksøyr J. Comparison of phenotypic diversity and DNA heterogeneity in a population of soil bacteria. 1990, *Applied and Environmental Microbiology*; 56: 776-81.
- Torsvik V., Øvreås L. Microbial diversity and function in soil: from genes to ecosystems. 2002, *Current Opinion in Microbiology*; 5: 240-5.
- Tsushima I., Kindaichi T., Okabe S. Quantification of anaerobic ammonium oxidizing bacteria in enrichment cultures by real-time PCR. 2007, *Water Research*; 41: 785-794.
- van de Graaf AA, Mulder A, de Bruijn P, Jetten MS, Robertson LA, Kuenen JG. Anaerobic oxidation of ammonium is a biologically mediated process. 1995, *Applied and Environmental Microbiology*; 61: 1246-51.
- Van de Peer Y., Chapelle S., De Wachter R. A quantitative map of nucleotide substitution rates in bacterial rRNA. 1996, *Nucleic Acids Research*; 24: 3381-91.
- Van der Weerden, Jarvis S.C. Ammonia emission factors for N fertilizers applied to two contrasting grassland soils. 1997, *Environmental Pollution*, 95, 205-211 .
- Van Gylswyk N.O., Van der toorn J.J.T.K. Description and designation of a neo type strain of *Eubacterium cellulosolvens* (*Cillobacterium cellulosolvens* Bryant, Small, Bouma, and Robinson) Holdeman and Moore. 1986, *International Journal of Systematic Bacteriology*; 36: 275-277.
- Van Niftrik, L., Fuerst, J.A., Sinninghe Damstè, J.S., Kuenen, J.G., Jetten, M.S.M., Strous, M., The anammoxosome: an intracytoplasmic compartment in anammox bacteria. 2004, *FEMS Microbiology Letters*; 233: 7 - 13 .
- Van Niftrik, L., van Helden, M., Kirchen, S., van Donselaar, E.G., Harhangi, H.R., Webb, R.I., Fuerst, J.A., den Camp, H.J.M.O., Jetten, M.S.M., Strous, M., Intracellular localization of membrane-bound ATPases in the compartmentalized anammox bacterium

- 'Candidatus *Kuenenia stuttgartiensis*'. 2010, *Molecular Microbiology*; 77: 701 - 715 .
- Van Soest, P.J. *Nutritional Ecology of the Ruminant*, second ed. 1994, Cornell University press, Ithaca, NY, pp244-245.
  - Vartoukian S.R., Palmer R.M., Wade W.G. Strategies for culture of 'unculturable' bacteria. *FEMS Microbiol Lett.* 2010 309: 1-7.
  - Vucebic S. The mechanism of gold extraction and copper precipitation from low grade ores in cyanide ammonia systems. 1997, *Minerals Engineering*, 10, 309-326 .
  - Von Engelhardt W., Dycker C.H., Lechner-Doll M. Absorption of short-chain fatty acids, sodium and water from the forestomach of camels. 2007, *Journal of Comparative Physiology B*; 177: 631–640
  - Wachi M., Umitsuki G., Shimizu M., Takada A., Nagai K. *Escherichia coli* cafA gene encodes a novel RNase, designated as RNase G, involved in processing of the 5' end of 16S rRNA. 1999, *Biochemical and Biophysical Research Communications*; 259: 483-8.
  - Wagner M, Horn M. The Planctomycetes, Verrucomicrobia, Chlamydiae and sister phyla comprise a superphylum with biotechnological and medical relevance. *Current Opinion in Biotechnology* 2006 ;17: 241-9.
  - Waki M, Tokutomi T, Yokoyama H, Tanaka Y. Nitrogen removal from animal waste treatment water by anammox enrichment. 2007, *Bioresource Technology*; 98: 2775-80.
  - Wayne L.G., Brenner D.J., Colwell R.R., Grimont P.A.D., Kandler O., Krichevsky M.I., Moore L.H., Moore W.E.C., Murray R.G.E., Stackenbrandt E., Starr M.P., Truper H.G. Report of the Ad Hoc Committee on Reconciliation of Approaches to bacterial systematics. 1987, *International Journal of Systematic Bacteriology*; 37: 463-364.
  - Wayne A., Fournier D., Fournier M.J. rRNA modifications and ribosome function. 2002, *TRENDS in Biochemical Science*, 27: 344-351.
  - Walsh D.A., Baptiste E., Kamekura M., Doolittle W.F. Evolution of the RNA polymerase B' subunit gene (rpoB') in Halobacteriales: a complementary molecular marker to the SSU rRNA gene. 2004, *Mol Biol Evol*; 21: 2340-51.
  - Wang H.C., Hickey D.A. Evidence for strong selective constraint acting on the nucleotide composition of 16S ribosomal RNA genes. 2002, *Nucleic Acids Research*; 30: 2501-7.
  - Wang G., Luo H., Meng K., Wang Y., Huang H., Shi P., Pan X., Yang P., Diao Q., Zhang H., Yao B. High genetic diversity and different distributions of glycosyl hydrolase family 10 and 11 xylanases in the goat rumen. 2011, *PLoS One*; 6: e16731.
  - Warnecke F, Luginbühl P, Ivanova N, Ghassemian M, Richardson TH, Stege JT, Cayouette M, McHardy AC, Djordjevic G, Aboushadi N, Sorek R, Tringe SG, Podar M, Martin HG, Kunin V, Dalevi D, Madejska J, Kirton E, Platt D, Szeto E, Salamov A, Barry K, Mikhailova N, Kyrpides NC, Matson EG, Ottosen EA, Zhang X, Hernández M, Murillo C, Acosta LG, Rigoutsos I, Tamayo G, Green BD, Chang C, Rubin EM, Mathur EJ, Robertson DE, Hugenholtz P, Leadbetter JR. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature*. 2007 ;450: 560-5.
  - Weimer P.J., French A.D., Calamari T.A.. Differential fermentation of cellulose allomorphs by ruminal cellulolytic bacteria. 1991, *Applied and Environmental Microbiology*; 57: 3101-3106.

- Weimer PJ, Odt CL. Cellulose degradation by ruminal microbes: physiological and hydrolytic diversity among ruminal cellulolytic bacteria. Saddler J, Penner M, editors. *Enzymatic Degradation of Insoluble Carbohydrates*. Washington, DC: American Chemical Society; 1996. pp. 291–304.
- Weimer P.J., Russell J.B, Muck R.E. Lessons from the cow: what the ruminant animal can teach us about consolidated bioprocessing of cellulosic biomass. 2009, *Bioresource Technology*; 100: 5323-31.
- Wherland S, Farver O, Pecht I. Intramolecular electron transfer in nitrite reductases. 2005, *ChemPhysChem*; 6: 805-12.
- Wilck M.B., Wu Y, Howe J.G, Crouch J.Y., Edberg S.,C. Endocarditis caused by culture-negative organisms visible by Brown and Brenn staining: utility of PCR and DNA sequencing for diagnosis. 2001, *Journal of Clinical Microbiology*; 39: 2025-7.
- Williams A.G., Withers S.E. *Bacillus spp.* in the rumen ecosystem. Hemicellulose depolymerases and glycoside hydrolases of *Bacillus spp.* and rumen isolates grown under anaerobic conditions. 2008, *Journal of Applied Microbiology*; 55: 283-292.
- Wimberly B.T., Brodersen D.E., Clemons W.M., Morgan-Warren R.J., Carter A.P., Vornrhein C., Hartsch T., Ramakrishnan V. Structure of the 30S ribosomal subunit. 2000, *Nature*: 407: 327-39.
- Woese C.R., Fox G.E., Zablen L., Uchida T., Bonen L., Pechman K., Lewis B.J., Stahl D. 1975 Conservation of primary structure in 16S ribosomal RNA. *Nature*; 254: 83-6.
- Woese C.R. Bacterial Evolution. 1987, *Microbiological Review*; 51: 221-271
- Wojciechowicz M., Ziółcki A. Pectinolytic enzymes of large rumen treponemes. 1979, *Applied and Environmental Microbiology*; 37: 136-42.
- Woodson S.A., Leontis N.B. Structure and dynamics of ribosomal RNA. 1998, *Current Opinion in Structural Biology*; 8: 294-300.
- Wooley J.C., Godzik A., Friedberg I. A primer on metagenomics. 2006, *PLoS Computational Biology*; 6: e1000667.
- Wuyts J., Van de Peer Y., De Wachter R. Distribution of substitution rates and location of insertion sites in the tertiary structure of ribosomal RNA. 2001, *Nucleic Acids Research*; 29: 5017-28.
- Y Kamagata, H Tamak Cultivation of uncultured fastidious microbes *Microbes and Environments* Volume: 20, Issue: 2, Publisher: J-STAGE, Pages: 85–91
- Yang S., Ma S., Chen J., Mao H., He Y., Xi D., Yang L., He T., Deng W. Bacterial diversity in the rumen of Gayals (*Bos frontalis*), Swamp buffaloes (*Bubalus bubalis*) and Holstein cow as revealed by cloned 16S rRNA gene sequences. 2010, *Molecular Biology Reports*; 37: 2063-73.
- Yoda K., Toyoda A., Mukoyama Y., Nakamura Y., Minato H. Cloning, sequencing, and expression of a Eubacterium cellulosolvens 5 gene encoding an endoglucanase with novel carbohydrate-binding modules, and properties of Cel5A. 2005, *Applied and Environmental Microbiology*; 71: 5787-93.

- Young R.A., Steitz J.A.. Complementary sequences 1700 nucleotides apart form a ribonuclease III cleavage site in *Escherichia coli* ribosomal precursor RNA. 1978, PNAS; 75: 3593-7.
- Yusupov M.M., Yusupova G.Z., Baucom A., Lieberman K., Earnest T.N., Cate J.H.D., Noller H. Crystal structure of the ribosome at 5.5Å resolution. 2001, Science; 292: 883-896.
- Zhang L., Zheng P., Tang C.J., Jin R.C. Anaerobic ammonium oxidation for treatment of ammonium-rich wastewaters. 2008, Journal of Zhejiang University SCIENCE B; 9: 416-26.
- Zhou J., Huang H., Meng K., Shi P., Wang Y., Luo H., Yang P., Bai Y., Zhou Z., Yao B. Molecular and biochemical characterization of a novel xylanase from the symbiotic *Sphingobacterium* sp. TN19. 2010, Applied Microbiology and Biotechnology; 85: 323-33.
- Ziölecki A., Wojciechowicz M. Small pectinolytic spirochetes from the rumen. 1980, Applied and Environmental Microbiology; 39: 919-22.
- Zumft W.G, Körner H. Enzyme diversity and mosaic gene organization in denitrification. 1997, Antonie Van Leeuwenhoek; 71: 43-58.
- Zumft W.G. Cell biology and molecular basis of denitrification. Microbiology and Molecular Biology Reviews 1997 ;61: 533-616.
- Zumft W.G. Nitric oxide reductases of prokaryotes with emphasis on the respiratory, heme-copper oxidase type. Journal of Inorganic Biochemistry. 99: 194-215.
- Zumft W.G. Biogenesis of the bacterial respiratory CuA, Cu-S enzyme nitrous oxide reductase. 2005, Journal of Molecular Microbiology and Biotechnology; 10: 154-66.