

# GENOME-CENTRIC METAGENOMIC INVESTIGATION OF 134 SAMPLES COLLECTED FROM BIOGAS REACTORS REVEALED IMPORTANT FUNCTIONAL ROLES FOR MICROBIAL SPECIES BELONGING TO UNDEREXPLORED TAXA

Campanaro, S.<sup>1</sup>, Luo, G.<sup>2</sup>, Treu, L.<sup>3</sup>, Kougias, P.G.<sup>4</sup>, Zhu X.<sup>3</sup>, Angelidaki, I.<sup>3</sup>

<sup>1</sup> *University of Padova, Department of Biology, Via U. Bassi, 58b, 35131 Padova (PD), Italy*

<sup>2</sup> *Fudan University, Department of Environmental Science and Engineering, Shanghai Key Laboratory of Atmospheric Particle Pollution and Prevention (LAP3), 200433, Shanghai, China*

<sup>3</sup> *Technical University of Denmark, Department of Environmental Engineering, Kgs. Lyngby DK-2800, Denmark*

<sup>4</sup> *Institute of Animal Science, Hellenic Agricultural Organization DEMETER, Paralimni, Greece*

e-mail of the corresponding author: stefano.campanaro@unipd.it

**Keywords:** anaerobic digestion; metagenomics; metagenome-assembled genomes; taxonomy.

## INTRODUCTION

In the past four years many samples collected from full-scale biogas plants and lab-scale reactors have been investigated using Illumina shotgun sequencing approaches and deposited in NCBI Sequence Read Archive database. In all these reactors, the Anaerobic Digestion (AD) process is performed by a plethora of different microorganisms organized in a complex functional network where different species have distinct roles in degradation of organic matter. Difficulties associated to cultivation of many prokaryotes present in natural and engineered ecosystems strongly limits the possibility to expand our knowledge regarding their physiology, genetics and function (Tringe and Rubin, 2005). Thanks to the bioinformatics approaches recently developed, which allow the reconstruction of microbial genomes starting from shotgun DNA sequences obtained from mixed cultures (Campanaro *et al.*, 2016), it is now becoming possible to reveal the functional roles of microbial species resisting cultivation in axenic cultures. This process, named “genome-centric” metagenomics, allows the reconstruction of the so-called Metagenome Assembled Genomes (MAGs) obtained through bioinformatics binning of scaffolds derived from metagenomic assembly. This “genome-centric” approach represents a novel powerful tool to investigate the phylogeny of uncultured microbes and to determine their metabolic potential. Therefore, the aim of this study is to clarify the functional role of species identified in 134 samples associated to the AD system, focusing attention on those associated to uncharted branches of the tree of life, previously referred as “microbial dark matter”.

## MATERIALS AND METHODS

The collected samples were obtained from full-scale biogas reactors located in different countries (including Denmark, Germany and Spain) and many laboratory-scale reactors operated in a wide range of different conditions. Shotgun Illumina sequences associated to these samples were downloaded from SRA database. Assembly and binning of shotgun reads were performed using procedures previously reported (Campanaro *et al.*, 2018). Quality of recovered MAGs and taxonomic assignment were performed using CheckM (Parks *et al.*, 2015) and PhloPhlAn (Segata *et al.*, 2013). A specific taxonomic investigation was performed on four selected MAGs (DTU308, DTU309, DTU310, DTU320) by selecting 5,278 representative genomes from NCBI microbial genomes database and building a tree using PhloPhlAn (Segata *et al.*, 2013). A manual taxonomic inspection supported by an interactive viewer for rooted phylogenetic trees (Dendroscope software) was implemented (Huson and Scornavacca, 2012). Functional analysis and metabolic reconstruction of the genomes was performed using RAST (Aziz *et al.*, 2008).

## RESULTS AND DISCUSSION

The genome-centric reconstruction performed starting from the 134 samples allowed the identification of 5194 MAGs. After a careful quality check performed to remove low-quality genomes, and a de-replication step targeted to remove genomes collected multiple times, 1635 MAGs were selected. After automatic taxonomic investigation performed using PhyloPhlAn software, numerous species still had uncertain taxonomy. Therefore, a detailed investigation was performed by selecting 5,278 representative genomes from the NCBI database, and building a comprehensive phylogenetic tree based on multiple sequence alignment of 400 taxonomic informative proteins. After manual verification, fifty-four MAGs were associated to the Candidate Phyla Radiation and three to the DPANN radiation of archaea. Other genomes were assigned to poorly classified taxa such as *Candidatus Atribacteria* and *Candidatus Fermentibacteria*. After taxonomic analysis and evaluation of relative abundance in different samples, four of these MAGs were selected for a more accurate functional analysis. MAG “DTU308” was related to the *Candidatus Falkowbacteria* group, “DTU309” to the *Candidatus Shapirobacteria*, “DTU310” to the *Candidatus Moranbacteria* and “DTU320” to the candidate division Hyd24-12.

These four MAGs were identified in biogas plants associated to the WWTP, suggesting that they have a specific functional role in digesting sludge derived from wastewater treatment. On contrary, their abundance was extremely low in mesophilic and thermophilic “manure-based” biogas reactors. DTU320 was the most abundant among the four MAGs, and in Maabjerg biogas plant it represented ~22% of the entire microbiome. The second most abundant MAGs was DTU308 which represented ~2.4% of the entire microbial community in Avedøre biogas plant. These findings indicate that some species assigned to underexplored taxa represent an important component of the AD microbiome.

Annotation of the proteins encoded from the genome sequences predicted possible functional roles of specific MAGs. DTU320 encodes numerous proteins involved in monosaccharides utilization and, particularly, in mannose metabolism, D-ribose utilization, deoxyribose/deoxynucleoside catabolism and fructose utilization. Moreover, the presence of a partial Wood-Ljungdahl pathway (including the formyltetrahydrofolate synthetase-*fhs* gene), suggests a role in acetate utilization. Additionally, a *rnf*-like group of electron transport proteins indicated the presence of energy-conservation mechanisms via the energy-converting ferredoxin: NAD<sup>+</sup> reductase complex. Some findings derived from the metabolic reconstruction of DTU310 highlighted the central role of pyruvate in different pathways of this MAG. Regarding DTU309 and DTU310, the presence of ABC transporters for oligopeptides, dipeptides and amino acids supports a potential involvement in protein utilization. Finally, in DTU308, the low genome completeness still limits a clear understanding of its putative functional role.

## CONCLUSION

These findings suggest that, in the AD microbiome, metagenomics is an extremely efficient method for characterizing unknown microbial species associated to candidate taxa. Functional prediction of the proteins is important to uncover the role of microbes in the AD food-chain, particularly regarding the utilization of crucial compounds such as acetate and proteins.

## REFERENCES

- Aziz, RK, Bartels, D, Best, AA, DeJongh, M, Disz, T, Edwards, RA, et al. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9, 75.
- Campanaro, S, Treu, L, Kougias, PG, De Francisci, D, Valle, G, Angelidaki, I. 2016. Metagenomic analysis and functional characterization of the biogas microbiome using high throughput shotgun sequencing and a novel binning strategy. *Biotechnol Biofuels* 9, 26.
- Campanaro, S, Treu, L, Kougias, PG, Luo, G, Angelidaki, I. 2018. Metagenomic binning reveals the functional roles of core abundant microorganisms in twelve full-scale biogas plants. *Water Res* 140, 123-134.
- Huson, DH, Scornavacca, C. 2012. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst Biol* 61: 1061-1067.
- Luo, G, Fotidis, IA, Angelidaki, I. 2016. Comparative analysis of taxonomic, functional, and metabolic patterns of microbiomes from 14 full-scale biogas reactors by metagenomic sequencing and radioisotopic

analysis. *Biotechnol Biofuels* 9, 51.

Parks, DH, Imelfort, M, Skennerton, CT, Hugenholtz, P, Tyson, GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25:1043-55.

Segata, N, Börnigen, D, Morgan, XC, Huttenhower, C. 2013. PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. *Nat Commun* 4, 2304.

Tringe, SG, Rubin, EM. 2005. Metagenomics: DNA sequencing of environmental samples. *Nat Rev Genet* 6, 805-814.

### **Acknowledgement**

The study was funded by National Key Research and Development Program of China (2015BAD15B06, 2017YFC0212205), National Natural Science Foundation of China (51408133, 51750110498), State Key Laboratory of Pollution Control and Resource Reuse Foundation (PCRRF16009), and Science and Technology Commission of Shanghai Municipality (17230740900).