Strain resolved metagenomics applied to biogas upgrading

Gabriele Ghiotto[1], Stefano Campanaro[1,3](*), Laura Treu[1]

[1] Department of Biology, University of Padua, via U. Bassi 58/b, 35131 Padova, Italy
[3] CRIBI biotechnology centre, University of Padua, via U. Bassi 58/b, 35131 Padova, Italy
Keywords: biogas upgrading, metagenomic, strain, variants
Presenting author email: gabriele.ghiotto@phd.unipd.it

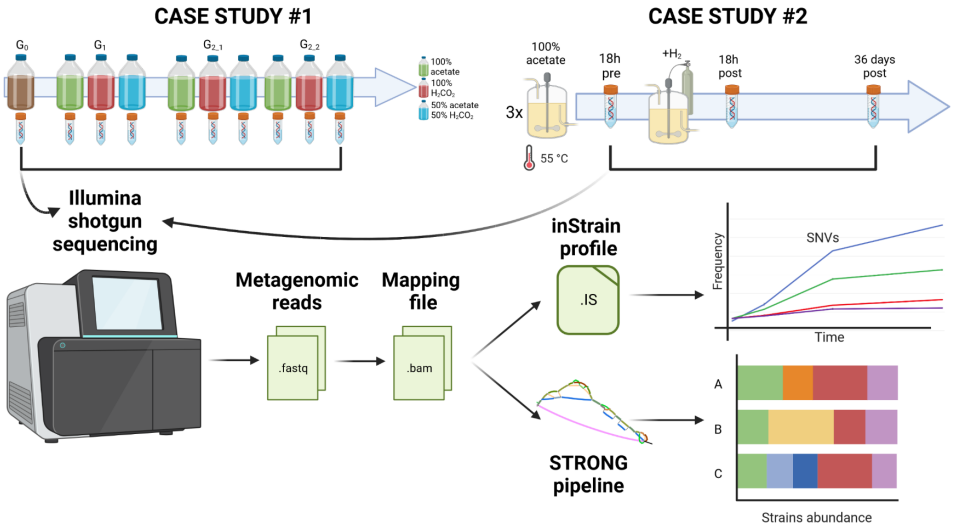

Figure 1. Graphical abstract representing the experimental setup of the two case studies and the bioinformatic pipeline used.

1. Introduction

Carbon Capture and Utilization (CCU) has gained significant attention in recent years as a crucial solution in mitigating greenhouse gas emissions and addressing climate change (Sabri et al., 2021). Among the various applications of CCU, the production of green biogas stands out as a promising renewable energy source. Green biogas upgrading, the process of removing impurities such as carbon dioxide (CO_2) and hydrogen sulfide (H_2S) from biogas to increase its energy content and reduce its corrosiveness, is critical for the wider adoption of biomethane as a replacement for traditional fossil fuels (Angelidaki et al., 2018). Methanogens are microorganisms that play a key role in the production of biogas through the anaerobic digestion of organic matter. These microbes are responsible for the conversion of CO_2 and hydrogen (H_2) into methane (CH_4). Hydrogenotrophic archaea are a specific group of methanogens that use hydrogen as an energy source (Lai et al., 2021). During the upgrading process, they can regulate H_2 levels and help to remove sulfur and acetate from biogas in a syntrophic relationship with sulfate-reducing bacteria (SRB) and sulfate-reducing acetogens (SAOB) (Zhu et al., 2017). The evolution of these microorganisms is driven by de novo mutations and changes in variant frequencies over time. Single Nucleotide Variants (SNVs) are a form of genetic variation that results from a single base pair change in the DNA sequence. These variations can have a significant impact on the metabolic pathways and functional properties of these microorganisms (Garud et al., 2019; Roodgar et al., 2021). The accumulation of SNVs can result from a variety of factors, including selective pressures applied to microbial communities during the biogas production process . In particular, the addition of specific feedstock substrates or changes in environmental conditions (pH, temperature, etc.) can drive the rise of new strains with distinct functional properties. These new strains may exhibit improved metabolic capabilities, which can lead to greater efficiency in the biogas production process. Studying the impact of SNVs in hydrogenotrophic archaea is a novel area of research with the potential to reveal new insights into the genetic heterogeneity of these microorganisms and its impact on biogas upgrading. In this study, we aim to investigate the impact of SNVs on the functional

properties of hydrogenotrophic archaea in green biogas upgrading. By testing different feedstock substrates and applying constant H_2 addition to the mixed microbial community, we will examine how the microorganisms respond to these conditions and unveil strain-level metabolic variation. Our research aims to provide a deeper understanding of the genetic heterogeneity of hydrogenotrophic archaea and its impact on biogas upgrading.

2. Materials and methods

2.1. Metagenomic analysis

A genome-centric metagenomics approach was used to analyze microbial genomes in both the experiments. Reads were filtered by Trimmomatic v0.39 (Bolger et al., 2014) to remove adapters and low-quality bases. Short-read co-assembly was performed with Megahit v1.29 (Li et al., 2015) and long-read co-assembly was performed using Flye v2.9 (Kolmogorov et al., 2019). Binning was performed with Concoct v1.1.0 (Alneberg et al., 2014), MaxBin v2.2.7 (Wu et al., 2016), MetaBAT2 v2.15 (Kang et al., 2019), and VAMB v3.0.2-1 (Nissen et al., 2021). Results were de-replicated and aggregated with dRep v3.4.0 (Olm et al., 2017) to obtain final metagenome-assembled genomes (MAGs). CheckM v1.2.1 (Parks et al., 2015) was used to assess MAG quality and determine relative abundance, and GTDB-Tk v2.1.0 (Chaumeil et al., 2020) was used for taxonomic classification. Prodigal v2.6.3 (Hyatt et al., 2010) was used for open reading frame prediction and functional classification was done with eggNOG-mapper v2.1.9 (Cantalapiedra et al., 2021). Identifiers were assigned to the MAGs based on taxonomic level, binning tool, and number.

2.2. Strain-level metagenomics

Variant analysis was performed using the software InStrain v1.6.3 (Olm et al., 2021) on the high quality MAGs of both experiments. The InStrain profile module takes a FASTA file of the panel of MAGs and a BAM file for each sample, along with a scaffold-to-bin file and gene annotation file. Three optional parameters, --min_mapq 2, --min_read_ani 0.98 and –min_genome_coverage 1, were used. A panel of MAGs was selected based on coverage and variant metrics, and strain deconvolution was performed with STRONG (Quince et al., 2021).

3. Results and discussion

3.1. Case study #2: testing different substrates

The microbial community consisted of 47 high-quality MAGs, belonging to four phyla: Firmicutes (94%), Euryarchaeota (2%), Chloroflexi (2%), and Tenericutes (2%). The only archaeon present was Methanothermobacter wolfeii, which had a similar number of SNVs among the three different substrates (acetate, H_2CO_2, and 50% acetate + 50% H_2CO_2). The frequency of SNVs over time showed that in all three conditions, a group of variants became fixed while another decreased in frequency. The abundance of the strains confirmed that at generation G0, one strain of M. wolfeii was dominant, but at generation G2, another one took over. Non-synonymous SNVs of the new strain were associated with genes including Mtr, Mcr, Frh, Hdr, and Fwd, all part of the methanogenesis pathway (Evans et al., 2019). This result demonstrated that SNVs can affect methanogens, enhancing their methanogenic potential through the selection of more efficient strains.

3.2. Case study #1: H_2 addition

The 50 high-quality MAGs obtained through the assembly and binning process were assigned to 13 phyla, with Firmicutes and Euryarcheota accounting for 50% and 12% of the microbial community, respectively. The dominant archaea were *Methanosarcina thermophila* and *Methanocullus thermophilus*. The first did not seem to be under any selective pressure, while the second one had a high number of SNVs. The comparison between the SNVs' trajectories over time and the results of the strain deconvolution showed that *M. thermophila* was not impacted by the H_2 addition due to its preference for acetate as a substrate. On the other hand, the H_2 addition caused *M. thermophilus* to shift towards a more hydrogenotrophic metabolism, with non-synonymous SNVs targeting key enzymes involved in methanogenesis becoming fixed at the final timepoint. Moreover, a strain replacement event occurred, resulting in a new strain becoming dominant due to its evolutionary advantage in terms of fitness.

4. Conclusions

In conclusion, studying SNVs in hydrogenotrophic archaea is critical in improving our understanding of metabolic changes in these microorganisms and advancing CCU towards a more sustainable future. Additionally, the biotechnological impact of this study will aid in optimizing biogas production and unlocking the full potential of green biogas as a sustainable energy source.

5. References

Alneberg, J., Bjarnason, B. S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., Lahti, L., Loman, N. J., Andersson, A. F., & Quince, C. (2014). Binning metagenomic contigs by coverage and composition. *Nature Methods*, *11*(11), Article 11. https://doi.org/10.1038/nmeth.3103

Angelidaki, I., Treu, L., Tsapekos, P., Luo, G., Campanaro, S., Wenzel, H., & Kougias, P. G. (2018). Biogas upgrading and utilization: Current status and perspectives. *Biotechnology Advances*, *36*(2), 452–466. https://doi.org/10.1016/j.biotechadv.2018.01.011

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P., & Huerta-Cepas, J. (2021). eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Molecular Biology and Evolution*, *38*(12), 5825–5829. https://doi.org/10.1093/molbev/msab293

Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., & Parks, D. H. (2020). GTDB-Tk: A toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*, *36*(6), 1925–1927. https://doi.org/10.1093/bioinformatics/btz848

Evans, P. N., Boyd, J. A., Leu, A. O., Woodcroft, B. J., Parks, D. H., Hugenholtz, P., & Tyson, G. W. (2019). An evolving view of methane metabolism in the Archaea. *Nature Reviews Microbiology*, *17*(4), Article 4. https://doi.org/10.1038/s41579-018-0136-7

Garud, N. R., Good, B. H., Hallatschek, O., & Pollard, K. S. (2019). Evolutionary dynamics of bacteria in the gut microbiome within and across hosts. *PLOS Biology*, *17*(1), e3000102. https://doi.org/10.1371/journal.pbio.3000102

Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, *11*(1), 119. https://doi.org/10.1186/1471-2105-11-119

Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., & Wang, Z. (2019). MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*, *7*, e7359. https://doi.org/10.7717/peerj.7359

Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology*, *37*(5), 540–546. https://doi.org/10.1038/s41587-019-0072-8

Lai, C.-Y., Zhou, L., Yuan, Z., & Guo, J. (2021). Hydrogen-driven microbial biogas upgrading: Advances, challenges and solutions. *Water Research*, *197*, 117120. https://doi.org/10.1016/j.watres.2021.117120

Li, D., Liu, C.-M., Luo, R., Sadakane, K., & Lam, T.-W. (2015). MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, *31*(10), 1674–1676. https://doi.org/10.1093/bioinformatics/btv033

Nissen, J. N., Johansen, J., Allesøe, R. L., Sønderby, C. K., Armenteros, J. J. A., Grønbech, C. H., Jensen, L. J., Nielsen, H. B., Petersen, T. N., Winther, O., & Rasmussen, S. (2021). Improved metagenome binning and assembly using deep variational autoencoders. *Nature Biotechnology*, *39*(5), 555–560. https://doi.org/10.1038/s41587-020-00777-4

Olm, M. R., Brown, C. T., Brooks, B., & Banfield, J. F. (2017). dRep: A tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *The ISME Journal*, *11*(12), 2864–2868. https://doi.org/10.1038/ismej.2017.126

Olm, M. R., Crits-Christoph, A., Bouma-Gregson, K., Firek, B. A., Morowitz, M. J., & Banfield, J. F. (2021). InStrain profiles population microdiversity from metagenomic data and sensitively detects shared microbial strains. *Nature Biotechnology*, *39*(6), Article 6. https://doi.org/10.1038/s41587-020-00797-0

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*, *25*(7), 1043–1055. https://doi.org/10.1101/gr.186072.114

Quince, C., Nurk, S., Raguideau, S., James, R., Soyer, O. S., Summers, J. K., Limasset, A., Eren, A. M., Chikhi, R., & Darling, A. E. (2021). STRONG: Metagenomics strain resolution on assembly graphs. *Genome Biology*, *22*(1), 214. https://doi.org/10.1186/s13059-021-02419-7

Roodgar, M., Good, B. H., Garud, N. R., Martis, S., Avula, M., Zhou, W., Lancaster, S. M., Lee, H., Babveyh, A., Nesamoney, S., Pollard, K. S., & Snyder, M. P. (2021). Longitudinal linked-read sequencing reveals ecological and evolutionary responses of a human gut microbiome during antibiotic treatment. *Genome*

Research, *31*(8), 1433–1446. https://doi.org/10.1101/gr.265058.120

Sabri, M. A., Al Jitan, S., Bahamon, D., Vega, L. F., & Palmisano, G. (2021). Current and future perspectives on catalytic-based integrated carbon capture and utilization. *Science of The Total Environment*, *790*, 148081. https://doi.org/10.1016/j.scitotenv.2021.148081

Wu, Y.-W., Simmons, B. A., & Singer, S. W. (2016). MaxBin 2.0: An automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*, *32*(4), 605–607. https://doi.org/10.1093/bioinformatics/btv638

Zhu, X., Kougias, P. G., Treu, L., Campanaro, S., & Angelidaki, I. (2017). Microbial community changes in methanogenic granules during the transition from mesophilic to thermophilic conditions. *Applied Microbiology and Biotechnology*, *101*(3), 1313–1322. https://doi.org/10.1007/s00253-016-8028-0