Anaerobic digestion of Green Willow Biomass, a novel lignocellulosic substrate, and the evaluation of H_2-induced stress in anaerobic bioreactors by genome-resolved metatranscriptomics

PhD dissertation

Balázs Kakuk

Author

Kornél L. Kovács

Zoltán Bagi

Supervisors

Doctoral School of Environmental Sciences

Department of Biotechnology, Faculty of Science and Informatics

University of Szeged

2022

Szeged

Contents

1. A list of used abbreviations 6

2. Introduction 8

2.1. General scope 8

2.2. Willow as a substrate for bioenergy production 11

2.3. Biogas production and the utilization of lignocellulosic substrates 13

2.4. Improving biogas value by upgrading 14

2.5. Power-to-gas concept 15

2.6. Metagenomics and metatranscriptomics as a means to investigate the anaerobic digestion process and methanogenesis in particular 17

2.7. Microbial community changes induced by H_2-addition 22

3. Aims 24

4. Materials and methods 25

4.1. A small plantation of SRCW 25

4.2. Batch anaerobic digestions 25

4.3. Gas chromatography (GC) analysis 26

4.4. Total solid, volatile solid measurement 26

4.5. C/N ratio determination 27

4.6. Fiber analysis 27

4.7. 16S rRNS gene amplicon sequencing 27

4.8. Bioinformatics methods for metataxonomic analysis 28

4.9. Energy potential and methane potential calculations 29

4.10. Continuously stirred tank reactors (CSTRs) 30

4.11. Determination of fermentation parameters 30

4.12. Sampling and H_2-injection 31
4.13. DNA isolation 31
4.14. RNA isolation and cDNA synthesis 32
4.15. Total community DNA and cDNA sequencing 32
4.16. Bioinformatic methods for metagenome and metatranscriptome analyses 33
4.17. Availability of data and material 35
5. Results 36
5.1. Characterization of GWB as a biogas substrate 36
5.1.1. Dependence of the fermentation efficiency parameters on willow genotypes, harvest time and chemical composition of leaf and stem biomass 36
5.1.2. Leaf to stem ratios and methane production in mixed fermentation 44
5.1.3. Microbial taxonomic composition analysis of the digesters 46
5.1.4. The advantages of using green willow biomass in comparison to other biogas feedstocks 51
5.2. H_2-induced stress response analyzed by combined metagenomics and metatranscriptomics 56
5.2.1. General overview of the genome-resolved metatranscriptomics pipeline 58
5.2.2. Metagenomic binning results and taxonomical annotation 59
5.2.3. Comparison of the MAGs with the Biogas Microbiome (BGMB) database 62
5.2.4. Abundance of the MAGs 65
5.2.5. Metatranscriptomic pathways analysis 68
5.2.6. Changes in the expression levels in methanogenesis genes 69
5.2.7. Interactions between methanogenesis and other metabolic processes in Archaea 72
5.2.8. Changes in gene expression levels in bacterial MAGs 76

6. Discussion 78
7. Conclusions 89
8. Acknowledgements 93
9. References 95
10. Summary 121
11. Summary in Hungarian (Összefoglaló) 124
12. Supplementary Material 127

1. A list of used abbreviations

AAI: Average Amino Acid Identity

AD: Anaerobic digestion

ADF: Acid Detergent Fiber

ADL: Acid Detergent Lignin

ANI: Average Nucleotide Identity

ASV: Absolute (exact) sequence variant

BBU: Biological biogas upgrading

BGMB: Biogas Microbiome

DEG: Differentially Expressed Gene

DGE: Differential Gene Expression

DPW: Differentially Expressed Pathway

EN: Energo cultivar of *Salix viminalis* (diploid genotype)

PP: PoliPlus cultivar of *Salix viminalis* generated from EN (tetraploid genotype)

E_p: extractable energy potential of a biomass via a certain process, e.g. biomethanation (AD) or biomass burning, in GJ

GC: Gas chromatograph

GHG: Green-house gas

GWB: Green Willow Biomass

H2-MTR: Metatranscriptome samples from the H_2-fed biogas culture

KEGG: Kyoto Encyclopedia of Genes and Genomes

KO: KEGG Orthology

MAG: Metagenome-assembled genome

MG: Metagenomics

MT: Metatranscriptomics

N2-MTR: Metatranscriptome samples from biogas culture

NDF: Neutral Detergent Fiber

NGS: Next-generation sequencing

OTU: Operational Taxonomic Units

P2G: Power-to-Gas

P2H_2: Power-to-hydrogen

P2CH_4: Power-to-Methane

P2bioCH_4: Power-to-Biomethane

PWE: Pathway Enrichment

SRC: Short-rotation coppice

TIC: Total inorganic carbon

VOA: Volatile organic acids

VS: (total) Volatile Solids

WGS: Whole Genome Shotgun sequencing

WWB: Woody willow biomass

WSD: Willow saw-dust

2. Introduction

2.1. General scope

The impact of civilization on the Earth biosphere and crust is pervasive and persistent. The extent of this is so severe that our home planet was suggested to have entered the Anthropocene epoch, a proposed new geologic time unit distinct from the Holocene and earlier epochs (AWG, 2019; Steffen et al., 2018). The appearance of manufactured materials in sediments (plastic, concrete, etc.), and the alterations in biogeochemical cycles (C, N and P) over the past century substantially affected even the stratigraphic signal (Waters et al., 2016). Moreover, the worldwide biotic changes, including species invasions and the increasing rate of species extinction became severe and widespread even in high-biodiversity biomes such as rainforests and coral reefs. The rate of extinction is the highest since the Cretaceous-Quaternary extinction event 65 MYA (that wiped off the dinosaurs from the face of the Earth) (Barnosky et al., 2011). As a result, it is suggested that the sixth mass extinction event, known as the Holocene extinction is already under way (Ceballos et al., 2015); besides humanity's most likely contribution to the earlier Quaternary extinction event as well (the extinction of the majority of the megafauna in Eurasia and the Americas). Climate change, environmental pollution, habitat loss, overpopulation and the extensive utilization of fossil fuels all contribute to these severe and complex global issues (Folke et al., 2021). A very high level of multinational cooperation is necessary to cope with them. International treaties, such as the Kyoto Protocol (1992), the Paris Agreement (2015) or the Glasgow Climate Pact (2021) have been adopted in order to do so, but their success is far from guaranteed.

One central issue is the increasing global energy demand. The share of fossil fuels must be decreased in order to achieve the treaty's goals, and to minimize the increase of global average temperature (Secretariat REN21, 2021). Although the share of renewable and nuclear sources increases steadily, they still contribute to only 20% of the global energy share (IEA, 2021). All renewable energy sources have their advantages and disadvantages

(Secretariat REN21, 2021), there is no winner for every situation; rather, their combined usage is required to provide GHG-mitigation. Also, further research and development is necessary to make renewables economically feasible alternatives to fossil sources.

Among the renewable sources, biomass is perhaps the most widely utilized and also, it is the one that have been used since the dawn of humanity, as our ability to tame fire have played a major role in becoming human (Parker et al., 2016). Nowadays, biomass can be utilized in several ways either to generate electricity and/or heat directly via combustion in power plants for example, or it can be used to produce alternative biofuels. These, including, but are not limited to biodiesel, bioethanol, biobutanol and biohydrogen, are then mainly used for transportation purposes. While all have merits each has its own particular characteristics from typical substrates to conversion processes, by-products, as well as their state of technical development and various research and development challenges (Müller-Langer et al., 2014). Biogas, the product of biomass conversion is perhaps the most versatile. It is the result of the anaerobic digestion process, carried out by a very complex microbial community, which shows substantial differences between fermentors, depending on operational parameters and substrate types (Campanaro et al., 2020). Most of the time, biogas is used to generate electricity directly from it via combined heat and power engines (CHP), but it can also be stored and transported easily and used on site or in vehicles (Weiland, 2010). Biogas is composed of mainly CO_2 and CH_4 (with traces of other gases such as H_2 and H_2S), thus, after appropriate upgrading, the resulting purified biomethane is chemically equivalent to natural gas, thus it can be exploited in essentially every way the fossil natural gas is utilized today, i.e. it can be injected into the gas grid as well (Scarlat et al., 2018). Therefore, biomethane holds very high potential, altough efficient upgrading processes are needed to be developed. No wonder the biogas industry is growing dynamically worldwide. However, this is accompanied with increasing concerns about supplying the anaerobic digestion (AD) reactors with energy plants cultivated on agricultural land (Kalamaras and Kotsopoulos, 2014) such as maize silage (Purdy et al., 2017; Wagner et al., 2018). This is because the cultivation costs and artificial

fertilizer requirements of maize silage and similar crops is relatively high, and fertile agricultural lands should be used for food and feed production instead.

However, biogas has an additional advantage over other biofuels: it can be produced efficiently not just from biomasses with high contents of easily digestible carbohydrates, such as simple sugars or starch only, like bioethanol can be, but from harder-to-degrade lignocellulosic materials and biowastes as well. Therefore, intensive search is in progress to switch biogas production towards second generation, alternative substrates that do not come into conflict with crops that could serve human or animal nutritional purposes (food vs. energy debate) (Brethauer and Studer, 2015; Clifton-Brown et al., 2019). These include organic wastes and byproducts and lignocellulose-based substrates and utilizing them is both an environmentally and economically attractive bioenergy production technology.

In addition, improving the efficiency of biogas upgrading also draws scientific and commercial attention. Particularly promising is the power-to-biomethane (P2bioCH4) technology, an innovative type of power-to-gas process (P2G), which combines the advantages of anaerobic digestion technology with the more extensively employed green electricity producing technologies. This is by converting electricity into biomethane via feeding the biogas reactors with H_2 produced from the electrolysis of water (Bagi et al., 2017). The behavior of the microbial community, in particular the methane producing Archaea, in this technology is critically important, as they convert the external H_2 to CH_4.

In the context of climate change, the generation of biogas as a renewable energy form has become popular and was intensively examined over the last few decades (Kougias and Angelidaki, 2018). In the course of my doctoral work, I aimed to contribute to these two major challenges facing the biogas industry: 1.) I utilized a novel lignocellulosic substrate, the young shrubs of the short-rotation coppice willow (Salix spp.), which I termed Green Willow Biomass (GWB) for biogas production; and 2.) I investigated the effects of hydrogen-injection stress in the frame of a biological biogas upgrading scenario on the

biogas producing microbial community via metagenomic and metatranscriptomic approaches.

2.2. Willow as a substrate for bioenergy production

Second-generation lignocellulosic substrates are increasing in popularity as alternative substrates for bioenergy production because they can be produced at significantly lower costs than first generation energy crops (Gissén et al., 2014). Lignocellulosic biomass, the complex of biopolymers, composed of lignin, cellulose, and hemicelluloses, accounts for the highest proportion of renewable terrestrial biomass accumulated on Earth. Due to the complex microbial community residing in anaerobic digesters that can degrade basically any type of biological material, biogas is particularly promising in utilizing these types of biomasses.

Moreover, these second-generation substrates do not come into contradiction with the crops cultivated for food/feed purposes, such as maize or soybean as they can be produced not just on prime farmlands, but on marginal lands as well (Valentine et al., 2012). Although frequently the lignocellulosic plant material is burned in biomass furnaces, many of them can be or could be utilized in biogas reactors; however, their fermentation efficiency may not be good enough, thus their economic feasibility hinders their widespread usage (Wagner et al., 2018). Extensive research is addressed to the optimization of promising second-generation lignocellulosic substrates, including for example switchgrass (*Panicum virgatum*) and miscanthus (*Miscanthus × giganteus*) (Stoof et al., 2015). Another such plant, the Shrub willow (*Salix* spp.) can be cultivated as a short-rotation coppice (SRC willow, SRCW) with environmental and rural development benefits (Clifton-Brown et al., 2019; Fischer et al., 2005; Volk et al., 2016) for the efficient production of biomass, which than can be utilized to produce bioenergy and bioproducts. These benefits are as follows:

a) willow can be cultivated on marginal lands (Valentine et al., 2012);
b) it is perennial, therefore in current practice its woody stems are harvested in every 2–4 years (after a cutback in the first year) without needing replantation up to 7–10 harvesting cycles, which makes the plantation exploitable for over 15–20 years;
c) it is capable of carbon sequestration after a single crop cycle (seven 3-year rotations) (Caputo et al., 2014);
d) the net energy ratio, i.e., the ratio of energy gained over energy invested is in the range of 18:1 to 43:1 (Caputo et al., 2014);
e) it is resistant to environmental effects and stresses, endures contaminated soils, making it suitable for phytoremediation (Dimitriou et al., 2012; Janssen et al., 2015);
f) it has positive effects on species diversity (Cunniff and Cerasuolo, 2011; Volk et al., 2016);
g) it can be coupled with N-removal, as willow plants reduce soil-water nitrate-nitrogen concentrations by 40–80% (Ferrarini et al., 2017; Ssegane et al., 2016);
h) its cultivation is considerably less labor-demanding and costs less than that of sensitive herbaceous plants, and less prone to yield fluctuations (Caputo et al., 2014; Eisenbies et al., 2014).

Moreover, novel variants are continuously being developed, the main breeding approach to improve willow wood biomass yields and quality relies on species breeding to capture hybrid vigor (Fabio et al., 2017; Serapiglia et al., 2014). In 2016, Dudits et al. managed to duplicate the chromosome number of the Energo (EN) cultivar via a colchicine treatment, resulting in autotetraploid genotypes named PoliPlus (PP) (Dudits et al., 2016). The multiplied genome size increased leaf size, shoot diameter and root system, as well as it improved photosynthetic activity and plant hormone composition. With yield increases of 20–40% anticipated from breeding and selection efforts for new willow varieties (Serapiglia et al., 2013; Sleight and Volk, 2016), all these advantages and developments urge the utilization of SRC willow for a sustainable bioenergy production. Other ways to promote SRC willow is to decrease production associated costs.

Life-cycle analyses showed that harvesting the woody biomass is a major contributor to the willow production system: i.) costs associated with it contribute to the final production costs in the largest part; ii.) it is the second largest input of primary fossil energy in the system (after commercial N fertilizer); and finally iii.) it accounts for about one third of the energy input as well (Caputo et al., 2014; Eisenbies et al., 2014). Reducing it by using a more easily harvestable biomass can make the production even more economically feasible.

2.3. Biogas production and the utilization of lignocellulosic substrates

SRCW is utilized for bioenergy production mainly in the US, Canada, UK, Sweden (Clifton-Brown et al., 2019; Fabio et al., 2017) and Poland (Stolarski et al., 2019). Other Eastern-European and former Soviet Union countries also have a high biological energy potential (Fischer et al., 2005), including SRC willow (also poplar and Miscanthus). For example, in Latvia, Hungary and Lithuania this potential is the highest: more than 175 GJ/person. Within a usual current scenario, the ~3-year-old 'woody' SRCW biomass is harvested, chipped and incinerated in biomass furnace plants to generate heat and/or electricity. Few attempts used woody willow as biogas substrate: willow saw-dust (WSD) was shown to be an appropriate AD substrate following various physico-chemical (Estevez et al., 2012; Horn et al., 2011; Jurado et al., 2013; Uellendahl et al., 2008) and biological pretreatments (Alexandropoulou et al., 2017). The primarily aim of pretreatments are to partially deconstruct lignocellulose (Patinvoh et al., 2017), and lower its lignin content (Mulat et al., 2018b; van der Lelie et al., 2012), as the anaerobic degradation of this component proceeds very slowly, if at all (Mulat et al., 2018a); and because it surrounds the cellulose, making its enzymatic degradation more difficult (Montgomery and

Bochmann, 2014). In order to achieve this, most pretreatment methods require specialized and expensive instrumentations and/or conditions, although these are not always feasible at industrial scales.

Although higher lignin content is desirable for combustion purposes it is not for biofuel production (Tubeileh et al., 2016). The amount of lignin correlated negatively with CH_4 yields (Herrmann et al., 2016) in most crop species. Due to the recalcitrant properties of lignocelluloses, the hydraulic retention time (HRT) is increased and thus a lower organic loading rate (OLR), and also a high amount of dilution water has to be used, which makes their fermentation economically less feasible. These reasons limit the utilization of willow saw-dust, woody willow biomass (WWB), willow wood chips, or other woody biomass in industrial biogas plants.

2.4. Improving biogas value by upgrading

As tough selecting an appropriate substrate is very important in biogas production, there are other challenges yet to be fully addressed. These include developing cost-effective pretreatment methods [mentioned above and reviewed in (Brodeur et al., 2011; Frigon et al., 2012; Hendriks and Zeeman, 2009)] and the efficient upgrading of biogas into biomethane. Traditional biogas upgrading approaches include scrubbing, pressure swing adsorption and membrane separation, among others. In addition, biological biogas upgrading (BBU) can present an alternative, as relatively small capital investment is required, as it can be readily installed to existing biogas plants and because no high temperatures or pressures are needed for its operation (Sun et al., 2015). Yet another alternative is the relatively new power-to-gas concept. In this scenario the external H_2 is added to the system in order to reduce the CO_2 into CH_4 and thus to increase the methane content of biogas (Ács et al., 2019; Szuhaj et al., 2016).

BBU can be achieved *in situ*, when the H_2 is injected directly into the anaerobic digester, or *ex situ*, when a separate fermentation unit is used that contains enriched cultures of hydrogenotrophic methanogens. Both approaches have advantages and disadvantages: generally, *ex situ* processes can generate biomethane at a higher rate (although this depends on various process parameters, such as temperature, pressure, geometry), while *in situ* BBU costs are less as there is no need for investment for a second reactor and it can be easily coupled with the existing fermentor. Perhaps even more importantly, *ex situ* can produce biogas with up to 98% CH_4, which is adequate for grid injection even, whereas with *in situ,* the produced biogas can hardly exceed 85% CH_4 (Bagi et al., 2017). In a recent mini-review about *in situ* biogas upgrading the importance of *Methanoculleus* under mesophilic conditions and *Methanothermobacter* under thermophilic conditions was pointed out (Szuhaj et al., 2021; Zhang et al., 2020).

2.5. Power-to-gas concept

Although biogas technology has many advantages, such as its capacity to generate biofuel from diverse types of biomasses, including cheap lignocellulosics (for example corn stover or willow) and a wide range of waste streams derived from the agro-food industry (which is a source of vast amounts of readily degradable organic materials, composed mainly of complex organic molecules) and from liquid or solid communal waste treatments, it can't replace the fossil-fuel based energy production by itself. Instead, in order to achieve this, several different renewable energy production technologies are needed to be coupled together. Since currently, solar and wind based electricity are the most preferred ones by governments and these are technologically and economically the most advanced technologies (Secretariat REN21, 2021), they should be considered. In case of Hungary, as a result of the recent restrictive legislative changes regarding wind energy, the areas of renewables most likely to develop are solar and geothermal (Simon and Deák, 2020). But they have challenges as well; one major challenge is how to store the produced electricity. The amount of fluctuating renewable electricity depends on environmental factors

(sunshine and wind), and thus it is produced in an intermittent fashion. Therefore, it is very hard to produce them when they are needed, even more so because energy consumption is also fluctuating daily and seasonally (Cozzi and Goodson, 2020).

A large number of pathways exists for the transformation of renewable electricity in some type of energy storage platforms. One possibility is the Power-to-Gas (P2G) systems wherein gaseous fuels are produced using excess renewable electricity via water electrolysis (Hashimoto, 1994). In the Power-to-hydrogen ($P2H_2$) system H_2 is produced in an environmentally friendly route as there are no harmful by-products produced. However, as currently the storage and transportation infrastructure for H_2 remains under development, its large-scale utilization as energy carrier is yet to come. Doubtlessly, in the future H_2 will play a significant role. The infrastructure for transporting, storing and combusting (bio)methane (e.g. natural gas) on the other hand is very well established and developed, thus. Therefore, there is a large potential in the Power-to-Methane ($P2CH_4$) system, wherein methane is produced via combining the H_2 with carbon dioxide (CO_2), because this route combines the advantages of the preexisting natural gas grid with state-of-the-art technologies for photovoltaic and wind energy producing technologies (by using H_2 as intermediate energy carrier) (Bailera et al., 2017). Moreover, as CO2 is recycled, the global warming would be mitigated in some extent (Hashimoto et al., 1999). A recent, even more innovative system is the Power-to-Biomethane ($P2bioCH_4$), wherein the H_2 resulting from the P2G route is combined with biogas to produce biomethane via BBU. As mentioned, biomethane is advantageous because it can be used in any application for which fossil fuel natural gas is utilized today (Bagi et al., 2017).

2.6. Metagenomics and metatranscriptomics as a means to investigate the anaerobic digestion process and methanogenesis in particular

H_2 is a key component of the AD process, H_2-coupled electron transfer has been verified as an important extracellular pathway of sharing reducing agents within the anaerobic environment, especially in microbial electrosynthesis systems (Cai et al., 2020). The effects of excess H_2 on the anaerobic digestion microflora and on the methanogens can be examined at DNA and/or RNA level to shed light on the biological background of the P2bioCH4 process.

Metagenomics roots back to the late 70's, when Carl Woese pioneered the use of 16S rRNA gene sequences for the phylogenetic analysis of prokaryotes. During his research he realized that Archaea form a separate clade from Bacteria and thus laid ground for the 3-domains tree of life (Woese and Fox, 1977). With this technique, researchers may get answer to the question of '*Who is there?*'. Since then, mainly due to the lowering of the costs of NGS, targeted metagenome sequencing or *metataxonomics* is readily used to assess the bacterial, archaeal and/or fungal microbial community compositions – depending on the primers used (Yarza et al., 2014). 16S rRNA gene sequencing became a widely used procedure to analyze the microbial (prokaryotic) community compositions of anaerobic digesters (Hassa et al., 2018).

While these methods are cost-effective, they have drawbacks as well. They are prone to biases originating from the polymerase chain-reaction (PCR) step (introduced by the amplification primers) and from the difference in 16S rRNA gene copy number in the bacteria (Jünemann et al., 2017). In addition, the protocols differ in the combination of the nine hypervariable regions (HRVs) that are analyzed (Johnson et al., 2019). This, along with the fact that numerous bioinformatics approaches, pipelines and databases exist (Prodan et al., 2020), makes the comparisons between different studies difficult (Pollock et al., 2018). A main difference in the bioinformatic workflows is that whether they are

OTU-based (Operational Taxonomic Units), or ASV-based (absolute sequence variants) (Callahan et al., 2016). DADA2's ASV method was shown to have benefits over OTU-based ones, as it combines the benefits of both closed-reference and *de novo* OTU clustering approaches: the generated ASVs are reusable across studies, reproducible in future datasets and are not limited by incomplete reference databases (Callahan et al., 2017).

Amplification-free WGS metagenomics, or metagenomics for short, is performed when the total community DNA is sequenced (or aimed to be sequenced, to be more exact). WGS offers an alternative to metataxonomics for the analysis of the microbiome. The term metagenomics was first used by Handelsman et al back in 1998: „… the collective genomes of soil microflora, which we term the metagenome of the soil" (Handelsman et al., 1998). Metagenomics, in contrast to metataxonomics has the potential to reveal information on full genomes of the subject microorganisms, providing invaluable knowledge about their coding capacities or to give answers to the question '*What are they doing there?*', as well. Exact quantitative information about the expression of these coding regions can be achieved by a technique termed metatranscriptomics via sequencing the total mRNA population. In the following section I will summarize the scientific literature regarding metagenomics and metatranscriptomics applied to the study of biogas producing communities.

The earliest insights into the biogas-producing communities using 454 pyrosequencing was carried out in 2008 (Schlüter et al., 2008), but later others soon followed as the technology became more widespread (Cho et al., 2013; Jaenicke et al., 2011; Li et al., 2013). More knowledge was gained about the taxonomic compositions and genetic functions of the AD microbiome as the 454 technology became superseded by the SOLiD (R. Wirth et al., 2012) and Illumina technologies (Li et al., 2014).

16S marker gene based metagenomics complemented with WGS metagenomics was used to investigate the differences between a wet and a dry fermenter operations modes (Stolze et al., 2015). This revealed high similarity at higher taxonomic level but for example the

archaeal sub-communities differed significantly. A similar combination of WGS and 16S metagenomics demonstrated a significant difference between the taxonomic profiles of digesters at the phylum level, however, several OTUs for these phyla could not unambiguously be described taxonomically due to the lack of suitable reference genomes, therefore their functional roles remained unclear. This again demonstrated the limitations of 16S rRNA gene based analyses and highlighted the importance of analyzing the complete metagenome to understand the full functional capacity of the AD microbial community (Jünemann et al., 2017).

As sequencing costs reduced drastically, higher coverages and metatranscriptomics became possible. Employing deep sequencing and a combination of metagenomics and metatranscriptomics deepened our understanding of the AD microbiome (Bremges et al., 2015). The reconstruction of whole genomes eventually became possible, as sophisticated bioinformatics software were developed. Since its first emergence in 2013-2014 (Albertsen et al., 2013; Alneberg et al., 2014; Nielsen et al., 2014), this process, termed metagenomic binning has become the standard procedure to analyze microbiomes. The process involves the clustering ('binning') of each contig from a co-assembly, based mainly on their coverage into genome bins and then refining the bins based on sequence-specific information (e.g. GC content or tetranucleotide frequencies).

The first application of this method in biogas research was carried out on a laboratory-scale thermophilic CSTR, wherein 106 microbial genomes were assembled from ~51 Gbp of sequencing data (Campanaro et al., 2016). The results supported a funnel concept in the AD microbiome, in which a progressive functional specialization occurs while the process reaches the final step, methanogenesis. The authors identified key microbial genomes and the enzymes they encode, which are involved in specific metabolic pathways, such as carbohydrates' utilization, fatty acids degradation, amino acids fermentation, and syntrophic acetate oxidation. Soon after, a production-scale fermenter (Stolze et al., 2016) was investigated and five high-quality genome bins were assembled, presenting dominant but mostly unknown species belonging to the taxa Fusobacteria, Spirochaetes, *Cloacimonetes* and *Thermotogae*.

This collection of genomes was later expanded (Treu et al., 2016b) while also demonstrating the existence of a core group of microorganisms. Additional studies followed with the aim of expanding our knowledge of the population genomes of the biogas microbiome in different digesters and fermentation conditions. In an anaerobic digester treating pig manure and meadow grass the spatial distribution of lignocellulose-degrading microbes (planktonic or attached to the surface of the plant material) was investigated (Kougias et al., 2018), and revealed that *Bacteroides* and *Firmicutes* follow diverse metabolic strategies for polysaccharide degradation. In mesophilic and thermophilic full-scale biogas fermenters treating manure and sludge from wastewater treatment plants, an in-depth analysis regarding the biochemical pathways (based on the KEGG) and carbohydrate-active enzymes of the MAGs, the biotic and abiotic factors affecting their abundance, and the associations between the species was carried out, revealing functional roles of the core group of microorganisms and shedding light on the syntrophic interactions between them (Campanaro et al., 2018).

Specifically, the differences in MAG compositions in a mesophilic versus thermophilic biogas upgrading reactor showed that both community was dominated by *Methanoculleus* species, but from different populations and carrying different functions (L Treu et al., 2018).

Recently a study was aimed to reconcile the metagenomes from 134 different published sources to assemble the *biogas microbiome*, as is the collection of population genomes (MAGs) present in the investigated anaerobic digesters (Campanaro et al., 2020). The reads from each study were then mapped back to the MAGs (this process is also called 'fragment recruitment') to estimate their abundance in the samples. This revealed a great plasticity of the AD ecosystem and the presence of multiple different microbial communities that have little to no overlap among them.

A combination of metagenome and metatranscriptome analyses was complemented by the characterization of bacterial and archaeal isolates in a thermophilic biogas plant (Maus et al., 2016), and the authors argued that a high abundance in the metagenome does not

necessarily indicate high transcriptional or metabolic activity, and vice versa. Also, the thermophilic biogas plant comprised a huge number previously unknown species. Treu et. al investigated the effect of adding long-chain fatty acids (LCFA) and found that genes of the genus *Syntrophomonas* were upregulated indicating their importance in LCFA degradation (Treu et al., 2016a).

The influence of process temperature was investigated by metatranscriptomics in another study (Lin et al., 2016). Herein the authors used a lab-scale biogas reactors fermenting swine manure at 25 to 55 °C and their results revealed that the methanogenesis was centralized under thermophilic conditions – the diversity of functional pathways at higher temperatures was reduced.

In another combined metagenomics and metatranscriptomic study the active microbial populations during cellulosic biomass fermentation were investigated (Jia et al., 2018). The relative abundance of the transcriptionally active microbes and the analysis of their pathways showed that while many population genomes (PGs, in this context equivalent to MAG) were reconstructed, only a small subset was highly transcribed and that the trophic roles of the cultures were similar, the bacterial populations performing each function were distinct between the enrichment conditions. There was also a significant difference between the relative abundances of the MG and the MTR samples, supporting that metagenomics alone is insufficient for a complete analysis of the AD microbiome.

Improving the binning process has been the major goal of numerous improved bioinformatic algorithms (Alneberg et al., 2014; Graham et al., 2017; Kang et al., 2015; Wu et al., 2016); with several procedures developed that seek to increase the quality of the resulting bins. As the extent of metagenome sequencing grew rapidly, numerous approaches and pipelines were developed that enabled the analysis microbiomes via genome-resolved metagenomics: the analysis of MAGs – their abundance and function (Eren et al., 2015; Uritskiy et al., 2018). Complete workflows for genome-resolved metatranscriptomics are scarce tough. Perhaps SqueezeMeta (Tamames and Puente-Sánchez, 2019) and IMP (Integrated Meta-omic Pipeline) (Narayanasamy et al., 2016) are

the two that are the closest to this goal. These tools can analyze the actual function that is carried out by a microbial system at a given time point at RNA-level by estimating the abundance of either functional domains or the assembled MAGs and their relative changes. Humann2 is another sophisticated metatranscriptomics tool, however, it is read-centric, which makes it fundamentally different in several ways, for example, it doesn't pair the reads, doesn't carry out assembly and metagenomic binning and it uses the MetaCyc database for functional profiling. These metatranscriptomics tools are reviewed in (Shakya et al., 2019). None of these tools offer, however, the possibility to analyze the gene expression differences (aka. fold-change) on the resolution of each MAG.

Taken together, genome-centric metagenomics can reveal the genetic potential of the microbes, but only when it is combined with metatranscriptomics the detailed analysis of the specific functions carried out by each MAGs becomes possible.

2.7. Microbial community changes induced by H_2-addition

Within a P2CH_4 scenario, bursts of H_2 is converted into CH_4 by the hydrogenotrophic methanogens (Szuhaj et al., 2016), which are essential players in the AD microbial community. The development of a stable P2CH_4 microflora strongly depends on environmental conditions and on the starter microbial composition. Various reactor designs, operational parameters and inocula have been tested making rigorous comparison of the results difficult. In line with these conclusions, the mesophilic AD methanogenic community of palm oil mill effluent with eventual addition of formate was predominated by members of the genus *Methanoculleus* (Woraruthai et al., 2020). Various inocula were compared for biomethane production at mesophilic conditions in batch fermentations. It was concluded that the abundance and activity of the genera *Methanosarcina* and *Methanoculleus* played key roles in methanogenesis of added H_2, while the authors also noted the regulatory role of the available CO_2/bicarbonate in the production of CH_4 and/or volatile fatty acids.

In a recent work the microbial community changes were followed under various operational conditions starting from two distinct inocula, i.e., wastewater sludge and plug-flow reactor operated with agricultural waste (Logroño et al., 2021). The study pointed out the importance of the history of the inoculum communities. In the waste water inoculated batch reactors the methanogenic genera *Methanobacterium* and *Methanothrix* predominated and upon H_2 feeding the genus *Methanobacterium* took over. In the plug-flow reactor, supplied with animal manure and ensilaged plant biomass, the initial abundance of genus *Methanothrix* diminished, and the methanogenic gap was filled in by members of the genera *Methanobacterium* and *Methanoculleus*. This study corroborated the previous observations (Barua and Dhar, 2017; Batstone et al., 2006) concerning the regulatory role of H_2 concentration and CO_2 depletion in the selection of hydrogenotrophic methanogens predominating the P2bioCH_4 community. In a thorough in situ syngas bioconversion study running two up-flow anaerobic sludge blanket reactors in sequence at mesophilic temperature (Xu et al., 2020), observed the predominance of the genus *Methanothrix* (formerly *Methanosaeta*). The reactors were continuously fed with varying glucose loads. *Methanothrix* species apparently cannot carry out hydrogenotrophic methanogenesis, therefore their predominance under these conditions can be rationalized by the combined effects of glucose and CO-rich syngas addition via carboxydotrophic methanogenesis (Rafrafi et al., 2020). In addition, the recognized capability of *Methanothrix* species to carry out direct electron transfer to drive CO_2 reduction could facilitate the *Methanothrix* predominance (Nobu et al., 2015; Rotaru et al., 2014). In a recent study, thermophilic biogas reactors were fed with H_2, and after 18 hours and 36 days MTR analyses were carried out to unveil the involvement of the individual MAGs in the global microbiome functions (Zhu et al., 2020). Another study, which was the first that investigated biogas reactor metatranscriptome dynamics following hydrogen injection used a one week-difference (Fontana et al., 2018). No work was conducted that aimed to characterize the short-term response (i.e. few hours or maybe even minutes) of the methanogenic culture to H_2-addition. Since the intervals of the bursts of H_2 resulting from the wind and/or solar power generators may be short, insights from such analyses may

prove valuable for the of P2bioCH_4 technology. Moreover, since the high time-resolution metatranscriptomics of a complex microbial community has also not been conducted before, these results may offer general knowledge about the short-term changes as well.

3. Aims

In the first part of my doctoral work, I inspected a hypothesized scenario that a possible, more favorable alternative for the utilization of SRC willow could be to harvest the biomass while it is still in its 'green' form, termed as green willow biomass (GWB), i.e., less than one year old shrub. In order to exploit this possibility: 1.) I carried out a small-scale field trial to evaluate the biomass yields of both the stems and leaves of the GWB plants; 2.) assessed their substrate characteristics and specific methane yields (SMY) and highest methane production rates (μmax); 3.) assembled fermentations utilizing both the stems and leaves and examined the microbial community residing in these batch digesters via 16S amplicon sequencing, 4.) using the data from 1.) and 2.) I estimated the per hectare methane and exploitable energy potential, 5.) and finally compared these values to the woody biomass's and maize silage's corresponding data.

In the second part of my doctoral work, I aimed to analyze the effect of H_2 induced stress on the biogas-producing microbial community and in particular, the methanogenic Archaea. This experiment modelled a possible biological biogas upgrading scenario, the P2BioCH$_{4,}$ wherein a hydrogen is injected to an un-adapted microbial community.

4. Materials and methods

4.1. A small plantation of SRCW

An experimental plantation of *Salix viminalis* diploid Energo (EN) cultivar and its tetraploid PoliPlus (PP) forms (Dudits et al., 2016) was established in Szeged, Hungary. 100 cuttings of EN and PP cuttings were planted in February 2017, and in the following February, they were cut back to about 20 cm above ground. The plantation consisted of 20 rows, with 5 plants in each row. The distance of the plants were 50 cm, resulting in a planting density of 20,000 plants per hectare. This plantation was harvested in equal numbers in three harvests during 2018: June, August and October; each time, one full row was harvested. An early harvest in May was used for biomass productivity estimation only. The collected biomass, termed green willow biomass (GWB), was used in the experiments. Immediately after each harvest, the leaves and stems were separated, weighed and sampled and cut into approx. 10 cm long pieces, and finally frozen and stored at -20 °C for subsequent use. Before use, the frozen pieces were minced into approx. 1 cm long pieces.

4.2. Batch anaerobic digestions

Triplicate batch digesters were assembled for each combination of substrates with a substrate-inoculum ratio (S/I) of 0.5, at reaction temperature set to 37 °C. Inoculum for the standard biochemical methane potential (BMP) tests came from the industrial CSTR biogas facility of Zöldforrás Ltd., (Szeged, Hungary), which was fed with a mixture of pig slurry and maize and sweet sorghum silage at 37 °C. The batch experiments were carried out in 160 mL glass reactors in triplicates. Substrate concentration, the amount of inoculum and diluting water were calculated according to VDI 4630 protocol (see Kakuk et al., 2017). The fermentation volume was 60 mL, leaving a headspace of 100 mL. The reactors were flushed with N_2 to ensure anaerobic conditions and were sealed with butyl rubber stoppers and aluminum caps. Negative controls were included, in which only the

inoculum was used, to determine the endogenous methane production, which was subtracted from the net methane produced. BMP values were calculated by dividing the total methane yields with the input substrate's volatile solid (VS) content (mL CH_4/g VS). The fermenters were not stirred but were shaken manually each day before the GC measurement. Experiments were carried out for 45 days, the gas volume and the content of methane were monitored every day during the first week, and then every second or third day thereafter, until the end of the experiment. The modified Gompertz equation (Tjørve and Tjørve, 2017) was fitted to the actual data by using in-house developed R scripts with the growth rates R package (Petzoldt, 2018). The estimated BMPs (K-value, mL CH_4/g VS) and maximal production rates (μmax, mL CH_4/g VS/day) were derived from the fitted models.

4.3. Gas chromatography (GC) analysis

The CH_4 content was measured daily with Agilent 6890 N gas chromatograph (Agilent Technologies, Santa Clara, United States). The GC was equipped with HP Molesive 5 Å column (length 30 m, I.D. 0.53 megabore, film 25 μm) and thermal conductivity detector. The carrier gas was argon 5.0 (Linde Group Hungary, Budapest, Hungary) set at a flow rate of 16.8 mL/min. Split injection mode was applied (0.2:1), the injector temperature was 150°C. After each sampling and measurement, the headspaces of the fermenters were flushed with nitrogen gas (Messer Group, Bad Soden, Germany) for 10 min.

4.4. Total solid, volatile solid measurement

To determine the dry matter content (total solid, TS), the plant materials were kept at 105°C until their weight became constant. The volatile solid (VS) content was determined by placing the dried residues in an incinerator at 550°C for 2 h.

4.5. C/N ratio determination

An Elementar Vario MAX CN (Elementar Group, Hanau, Germany) analyzer was used to determine the C/N ratio of the substrates. The temperature of the combustion and post combustion tube was set at 900°C, the temperature of the reduction tube at 830°C. After the samples were burnt in the combustion tube the water vapor was separated by a specific adsorption column containing Sicapent (Merck KgaA, Darmstadt, Germany). The components were detected with a thermal conductivity detector. Helium was the carrier and flushing gas (Messer Group, Bad Soden, Germany).

4.6. Fiber analysis

The fiber composition determination, including the Neutral Detergent Fiber (NDF, Acid Detergent Fiber (ADF) and Acid Detergent Lignin (ADL) of the biomass samples were carried out with a VELP Scientific FIWE 3 Fiber Analyzer (VELP Scientifica Srl, Usmate, Italy) according to the manufacturer's guidelines using the Van Soest method (Goering and Van Soest, 1970). Hemicellulose was estimated as NDF-ADF, while cellulose as ADF-ADL. Lignin content was estimated as the ADL fraction.

4.7. 16S rRNS gene amplicon sequencing

In order to determine the microbial compositions of the digesters, a 150 mg sample from two, randomly selected reactors from each group of triplicate reactors was collected for total-community genomic DNA (gDNA) purification on day 30. After collection, the samples were stored at –20°C for further use. The extractions were performed with Quick-DNA Fecal/Soil Microbe Kit (Zymo Research Corporation, Irvine, USA) according to the manufacturer's instructions. DNA quantity was determined with Qubit 4.0 fluorimeter (Invitrogen, Waltham, USA) and DNA integrity was examined on 1% agarose gel. The V3-V4 region of 16S rDNA was amplified by PCR according to the Illumina protocol (Illumina,

San Diego, USA). The following primer pairs containing overhanging sequences were applied as recommended by the Manufacturer.

Forward Primer:

5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG-3′,

Reverse Primer:

5′-TCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC-3′.

A PCR reaction mixture (25 μL) containing 12.5 ng genomic DNA, 2x KAPA HiFi HotStart Ready Mix, 0.2 μM of each primer. The PCR reaction parameters were the following: initialization at 95°C for 3 min, followed by 25 cycles of denaturation, annealing and extension at 95°C, 65°C and 72°C for 30–30 s, respectively. Final elongation was performed at 72°C for 5 min. The libraries obtained were purified and then sequenced on an Illumina MiSeq platform. Illumina MiSeq® Reagent Kit v3 (2x300 bp) was used for the sequencing.

4.8. Bioinformatics methods for metataxonomic analysis

The read count data were treated according to the DADA2 pipeline (Callahan et al., 2016), which implements ASV (absolute sequence variant) identification that has a number of advantages over I-based (operational taxonomy unit) methods (Callahan et al., 2017). Read trimming and quality filtering was done with the DADA2's *filterAndTrim* function [parameters: truncLen = c(240,220), maxN = 0, maxEE = c (2,2), truncQ = 2, rm.phix = T, ctrimLeft = c(50, 55)], subsequently, *learnErrors*, *derepFastq*, dada and *removeBimeraDenovo* functions were used with default arguments. *assignTaxonomy* and *addSpecies* functions was used with SILVA-db, version 132, to annotate the reads (Quast et al., 2013).

The resulting sample/read count matrix was normalized via DESeq2's (Love et al., 2014) log-ratio transformation and subsequently the differentially abundant taxa between the

harvests and their log2 fold changes (*log2FC*) were assessed with the standard workflow (test = "Wald", fitType = „parametric), which uses the Wald test on fitted Negative Binomial GLMs (generalized linear models). This approach was shown to be an appropriate method for the gene abundance analysis of metagenomic samples (Jonsson et al., 2016). The significance threshold was set to p-value ≤ 0.05. Principial Component Analysis (PCA) was also carried out and for that the count matrix was transformed with DESeq's regularized log (rlog) transformation, which according to the package vignette's guidelines, has a variance stabilizing effect as this is more suitable for unsupervised clustering approaches (DESeq2 manual). The PCA results were visualized with the *factoextra* package (Kassambara et al., 2017).

4.9. Energy potential and methane potential calculations

To estimate the total VS yield of the plantation (t/ha), the VS yields of the measured plants were added up and normalized to the average planting density that was used (20,000 plant/ha). To estimate the average energy potentials via biogasification (MJ/plant), as the theoretical amount of energy that can be extracted from the plant biomass via AD, first the wet stem and leaf biomass values were multiplied with their respective VS % values and VS specific CH_4 yields to get the individual plant's CH_4 potential (CH_4/plant). Then the energy potential (MJ/plant) of the plants was estimated by multiplying this with the energy content (55.6 MJ/kg) and density (0.657 g/m^3, *https://encyclopedia.airliquide.com/methane?GasID=41*) of CH_4. To estimate the plantation's overall methane potential (CH_4 potential, ~ m^3/plantation) and energy potential (GJ/ha), the plant's CH_4 potential and energy potential were normalized to the planting density, respectively. In the case of WWB, the theoretical energy potential via combustion values were estimated by multiplying the combustion calorific values (CCVs) (measured with a calorimeter, see above) with the biomass yields (in terms of VS).

4.10. Continuously stirred tank reactors (CSTRs)

The anaerobic digestions were carried out in continuously stirred tank reactors (Kovács et al., 2013) with a fermentation volume of 5,000 mL and a headspace of 2,000 mL. The apparatus can be fed continuously or intermittently via a piston type delivery system, the fermentation effluent is removed through an air-tight overflow. The reactors are stirred with a spiral strip mixing device. Heated jacket surrounds the cylindrical stainless steel body, electrodes for the measurement of pH and redox potential are inserted through the reactor wall, in sealed sockets. The device can be drained at the bottom. The evolved gas leaves the reactor through the top plate, where ports for gas sampling and the delivery of liquids by means of syringes through silicone rubber septa are also installed. Gas volumes are measured with thermal mass flow devices (DMFC SLA5860S, Brooks). A fresh sample from an industrial scale mesophilic biogas plant, fed with pig slurry and maize silage mix (Zöldforrás Biogas Plant, Szeged, Hungary) was used as an inoculum, i.e., the microbial community adopted to heterogeneous substrate degradation. The reactors were flushed with N_2 to ensure anaerobic conditions and were closed airtight. Feeding regimes of the digesters were fed twice a day with synthetic medium in which only alpha-cellulose was added as a carbon source at a loading rate of 1 g oDM/L/day. The reactors were operated under mesophilic conditions, at 37 °C

4.11. Determination of fermentation parameters

Organic dry matter (ODM): The dry matter content was determined by drying the biomass at 105 °C overnight and weighing the residue. Further, heating of this residue at 550 °C provided the organic total solids content.

NH_{4+}–N: For the determination of NH_{4+}–N content, the Spectroquant Ammonium test (1.00683.0001 test, Merck, Kenilworth, N.J, USA) was used in a Nova 60 spectrophotometer according to the manufacturer's instructions.

VOA/TIC (Volatile organic acids/Total inorganic carbon): 5 g of each AD samples were taken for analysis and diluted appropriately with distilled water. The subsequent measurement was carried out with a Pronova FOS/TAC 2000 Version 812-09.2008 automatic titrator (Pronova, Germany).

4.12. Sampling and H_2-injection

The first set of samples were taken when the reactor operation was stabilized under N_2 in the headspace, the daily biogas production, CH_4 content and total organic acid/buffer capacity ratio were constant. 2 mL of reactor content was withdrawn and total RNA for transcriptome analysis (sample names: N2-MTR) and DNA for metagenome analysis (sample names: N2-MG) were isolated immediately after sampling. Then, the digesters were flushed with pure H_2 gas for 10 min on day 15 and 71. H_2 was injected directly from a pure H_2 gas cylinder through custom made nozzles (10 pieces) having 0.2 mm holes. The applied gas pressure was 2 bar, the gas purity was 99.999%. Two hours after flushing the reactors with H_2 samples were also taken for RNA (sample names: H2-MTR) and DNA (sample names: H2-MG) isolation. The headspace was then replaced with N_2 and the reactors were run under the same conditions as before. After 2 months the whole H_2 treatment procedure was repeated in order to test the reproducibility of the set-up. At the sampling time points two biological parallels were withdrawn.

4.13. DNA isolation

DNA extractions were carried out from 2 mL reactor liquid using a slightly modified version of the Zymo Research Fecal DNA kit (D6010, Zymo Research, Irvine, USA). The lysis mixture contained 100 µL of 10% CTAB (cetyltrimethylammonium bromide) to improve the efficiency (Wirth et al., 2015). After lysis (bead beating was performed by Vortex Genie 2, bead size: 0.1 mm; beating time: 15 min, beating speed: max) and subsequently the Zymo Research kit protocol was followed. The quantity of DNA was

determined in a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, United States) and a Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA, United States). DNA purity was tested by agarose gel electrophoresis and on an Agilent 2200 TapeStation instrument (Agilent Technologies, Santa Clara, CA, United States).

4.14. RNA isolation and cDNA synthesis

For RNA isolation 2 mL of reactor liquid samples were used. The samples were centrifuged at 12,000 rpm for 10 min. RNA extractions were carried out with the Zymo Research Soil/Fecal RNA kit (R2040, Zymo Research, Irvine, CA, United States). After lysis (bead beating), the Zymo Research kit protocol was followed. The DNA contamination was removed by Thermo Scientific RapidOut™ DNA removal kit (K2981, Thermo Fisher Scientific, Waltham, MA, United States). Ribosomal RNA was depleted using the Ribo-Zero™ rRNA Removal Kit for Bacteria (Illumina, Madison, USA) according to the manufacturer's instructions. The rRNA depleted samples were purified via the RNA Clean and Concentrator Columns from Zymo Research (Irvine, USA). During this step, an additional in-column DNase I treatment was included to ensure complete removal of DNA. Subsequently, synthesis of double-stranded cDNA was conducted using the Maxima H Minus Double-Stranded cDNA Synthesis Kit from ThermoScientific (Waltham, USA). In the first strand cDNA synthesis reaction, 2 μL of random hexamer primer were used. Final purification of the blunt-end double-stranded cDNA was carried out using SureClean Plus solution from Bioline (Luckenwalde, Germany). The quality of the RNA preparation was checked by agarose gel electrophoresis (data not shown).

4.15. Total community DNA and cDNA sequencing

The cDNA was sequenced in the same way as the total DNA. Paired-end libraries were prepared for the metagenome and metatranscriptome samples using the NEBNext® Ultra™

II DNA Library Prep Kit for Illumina (Cat.Num.: E7645L). Paired-end fragment reads were generated on an Illumina NextSeq sequencer using TG NextSeq® 500/550 High Output Kit v2 (300 cycles).

4.16. Bioinformatic methods for metagenome and metatranscriptome analyses

Quality filtering and trimming of the raw reads were carried out with FastQC (Andrews et al., 2010); metagenomic co-assembly from the filtered reads with MegaHIT (D. Li et al., 2016), ORF prediction on the assembled contigs with Prodigal (Hyatt et al., 2010) within the SqueezeMeta workflow (Tamames and Puente-Sánchez, 2019). The functional annotation of the predicted ORFs was carried out with EggNOGmapper (Huerta-Cepas et al., 2017) and for the KEGG KO annotation the EggNOG database (v. 5) was used (Huerta-Cepas et al., 2019). Binning of the contigs was carried out with four different binning algorithms: MetaBAT (Kang et al., 2015), MaxBin2 (Wu et al., 2016), Concoct (Alneberg et al., 2014) and Binsanity (Graham et al., 2017). The result of each binning procedure was further improved with DAS tool (Sieber et al., 2018). Bin qualities (completeness and redundancy) were estimated with CheckM (Parks et al., 2015), based on the presence/absence and number of single-copy marker genes. The taxonomy of the bins was determined using the Genome Taxonomy database (GTDB) using GTDB-Tk (Chaumeil et al., 2020). A phylogenomic tree from the protein genomes of the MAGs were built with the phylophlan3 program (Asnicar et al., 2020): *phylophlan –diversity high –fast -f phylophlan_configs/supermatrix_aa.cfg -t a –min_num_markers 75*. The assembly, annotation, binning and phylogenomic results were imported into and subsequently visualized with the Anvi'o platform (Eren et al., 2015). The results were compared with the Bio-Gas Microbiome database (Supplementary Table S2).

The filtered reads from each sample were mapped back onto each bins with bowtie2 (Langmead and Salzberg, 2012) and FeatureCounts (Liao et al., 2019) was used to calculate the gene count table by using the ORF predictions of the bins. Since we were

primarily interested in pathway analysis, genes that could be annotated with a KEGG Orthology (KO) were kept (Kanehisa et al., 2016).

The downstream data analysis was carried out within the R environment (R Core Team, 2019), with in-house developed scripts, relying heavily on the packages of the *tidyverse* (Wickham et al., 2019). First, the $\log_2$ fold changes (log2FC) between the control (N2) and treated (H2) samples were assessed using the DESeq2 package (Love et al., 2014). This was proven to be an appropriate method to infer differences between metagenomic and metatranscriptomic gene counts (Jonsson et al., 2016). The following parameters were set: test="Wald", fitType="parametric", filterFun = "ihw". For the assessment of significance, the Benjamini-Hochberg-adjusted p-values were used (termed '*padj*'), with a threshold of 0.05.

I assessed the differentially expressed KOs and pathways at two levels: First, counts of genes with the same KO annotation were grouped together and summed in each sample. Differentially expressed KOs between the two MTR samples were then determined with *DESeq2* as described above. The resulting DE KO list was the input for *Clusterprofiler* R package to detect differentially expressed pathways (Yu et al., 2012). Then counts of genes with the same KO annotation were grouped together in each sample and in each bin, since our main focus was to assess the pathways, which changed in the individual genome bins. Genes that did not belong to any bin were grouped together as *unbinned*. Differentially expressed KOs of every bin between the two MTR samples were then determined with *DESeq2*, based on *log2FC* and p-values. This bin-KEGGKO-sample table was also rlog-transformed (regularized logarithm transformation) with the *rlog* function of the *DESeq2* package and results were subjected to a Principial Component Analysis (PCA) using the *FactoMineR* package (Lê et al., 2008).

4.17. Availability of data and material

All the R scripts that were used to analyze the data are available upon request.

Raw read sequences for the 16S metataxonomic analyses of the batch reactors fed with GWB (*.fastq* files) are available on NCBI-SRA under the following BioProject id: *PRJNA701827* (https://www.ncbi. nlm.nih.gov/bioproject/PRJNA701827).

Raw read sequences for the WGS (N2_MG and H2_MG) and for the RNA-Seq data (N2_MTR and H2_MTR) (*.fastq* files) are available on NCBI-SRA under the following BioProject ID: PRJNA698464 (https://www.ncbi. nlm.nih.gov/bioproject/PRJNA698464).

5. Results

5.1. Characterization of GWB as a biogas substrate

The main goal of the first series of experiments was to compare the GWB substrate from the two different genotypes (EN and PP) harvested in three occasions, to the WWB substrate and to maize silage in terms of bioenergy potential. This is the net energy that can be extracted via biogasification from a unit area of plantation. These values can be estimated from the biomass yields and substrate specific methane yields (K), thus these parameters were measured as the first step. I tested both the leaves and stems, separately and simultaneously as well, using a leaf to stem ratio, in terms of VS, that was the same as the plants themselves showed. As reference, I also measured the characteristics of the woody willow biomass (WWB) on one- and two- years old woody stems harvested in wintertime. I measured chemical parameters affecting the efficiency of the substrate's anaerobic digestion (fiber composition), parameters that describe methane production rates (μmax), and microbial compositions that elucidate the process's microbiological background as well. Furthermore, I investigated the correlations between these parameters.

5.1.1. Dependence of the fermentation efficiency parameters on willow genotypes, harvest time and chemical composition of leaf and stem biomass

I harvested the small plantation of *Salix viminalis* Energo (EN) and PoliPlus (PP) consisting of 100 plants in equal numbers in three harvests during 2018 June, August and October. I The separated the leaves and stems and measured the biomass yields. Samples were taken and then the substrate characteristics (fiber composition, dry matter content, organic content) were assessed.

In my first series of experiments, I assembled batch anaerobic digesters from the diploid and tetraploid genotypes of GWB (and the WWB samples harvested in wintertime) and measured the methane yields via GC on a daily basis (later, after 20 days, the time between measurements were increased to 2-5 days). Subsequently I fitted the modified Gompertz model to the measured data. This growth model is commonly used in the scientific literature related to biogas production modelling to infer the properties of the dynamics of the biogas (or biomethane) formation, as it fits to the experimental values well (Zhao et al., 2017). From this model the *K* values, i.e. the maximal methane yields (mL CH_4/g VS), and the *μmax* values, i.e. the highest methane production rates (mL CH_4/day/g VS) can be estimated. These parameters describe the substrate's behavior in the bioreactors; and since both the speed of the methane production and the overall methane yields are important, especially from an industrial-scale fermentation point of view, they are of key importance. The results of the measurements are presented in **Table 1**.

The K-values of GWB leaves were on average 12% and 14% higher compared to GWB stems and to WWB (mean K-value = 232.5, 207.6 and 203.5 mL CH_4/g VS), respectively, but the difference was not statistically significant (*p-value* = 0.16). Leaves from both genotypes and stems from PP plants harvested in June served as superior CH_4 sources relative to the woody tissues, based on K-values. These differences were reduced in later vegetation phases, and then diminished in the samples collected in October. In contrast, the maximal CH_4 yield from the tetraploid stems harvested in this month was higher (194.3 mL CH4/g VS) than that of the diploid stems (148.7 mL CH4/g VS). Previously it was shown that the tetraploid plants developed thicker stems and thicker bark region and also wood formation between the primary and secondary xylem rings increased significantly (Dudits et al., 2016).

The *μmax* values showed significant difference between the GWB stems, leaves and WWB. Leaves showed generally faster CH_4 production rates (mean *μmax* = 17.5, 8.6 and 7.7 mL CH_4/day/g VS for GWB stems, leaves and WWB, respectively), which was 93% higher than GWB stems and 117% higher than WWB, on average (p-value < 0.0001). It is not

very surprising that green leaves are decomposed faster than stems, rather these results highlight the overall faster decomposition rate of GWB compared to WWB.

Harvest	Genotype	Plant	Organ	K	μmax	Lignin (ADL)	Hemicellulose (NDF-ADF)	Cellulose (ADF-ADL)	Solubles	C/N ratio
JUNE	EN	GWB	Stem	199.1	10.8	11.7*	15.8	32.5	40.1*	47.1
JUNE	PP	GWB	Stem	251.8*	11.2	8.9	14.9	40.2	36.0	43.2
JUNE	EN	GWB	Leaf	336.2	25.9	5.7	9.6	23.0	61.7	12.6
JUNE	PP	GWB	Leaf	334.9	24.8	7.3*	9.0	20.4	63.3	15.2
AUG	EN	GWB	Stem	221.7	8.6	21.6	13.2	38.6	26.7	57.2
AUG	PP	GWB	Stem	217.8	8.3	21.8	14.6*	38.5	25.1	54.0
AUG	EN	GWB	Leaf	186.6	14.1	12.8	12.5	21.6	53.1	21.5
AUG	PP	GWB	Leaf	199.8*	12.6	12.3	11.8	25.8	50.1	19.9
OCT	EN	GWB	Stem	148.7	5.4	27.6*	15.3*	32.4	24.8	70.0
OCT	PP	GWB	Stem	194.3	7.7*	23.1	13.5	34.6	28.8*	60.7
OCT	EN	GWB	Leaf	206.3*	14.7*	14.5	12.0	27.3	46.2	21.9
OCT	PP	GWB	Leaf	190.0	13.1	15.5	12.0	27.6	45.0	21.1
1-yo	EN	WWB	Stem	211.9	7.3	23.8	10.5	41.9	23.9	70.6
1-yo	PP	WWB	Stem	206.8	8.0	26.2	12.1	34.9	26.7	70.1
2-yo	EN	WWB	Stem	194.9	6.9	28.1	9.9	37.4	24.5	77.7
2-yo	PP	WWB	Stem	195.4	7.4	28.1	10.1	35.9	25.9	86.4
Mean standard deviation				11.91	0.79	0.65	0.44	1.48	1.04	2.31

Table 1. Fermentation and substrate parameters of the Energo (EN) and PoliPlus (PP) willow samples: green stems and leaves of each harvest in case of GWB; one- and two- years-old woody stems in case of WWB. Lignin, hemicellulose, cellulose and solubles content were measured according to the Van Soest method and are expressed as % dry mass. C/N ratio equals the measure of the substrates' total C content (% dry mass) divided by its total N content (% dry mass). K and *μmax* values mean the maximal CH_4 yields (mL CH_4/g VS) and highest CH_4 production rates (mL CH_4/day/g VS), derived from the modified Gompertz model, which was fitted to the CH_4 production data of each batch fermentation. The fermentations were carried out in triplicates, while the chemical analyses in duplicates. The mean standard deviation for each parameter is

indicated in the bottom row. In each bracketed sample group [according to harvest and plant part (Stem or Leaf)], an ANOVA was carried out to infer the difference between the mean values of EN and PP. A star (*) indicates if the mean value of a genotype was significantly higher than the mean of the other genotype in a respective comparison, under a *p-value* threshold of 0.05.

I carried out several statistical tests to assess the impact of the two factors in this first experimental arrangement, i.e. the harvesting time ('Harvest' factor: June, Aug, Oct) and genotype ('Genotype' factor: diploid EN and tetraploid PP) on the response variables, i.e. the K and μmax values. The tests were carried out separately for the stems and leaves.

On the one hand the effect of the Harvest and Genotype factors on the response variables were assessed simultaneously with a two-way ANOVA. Regarding the K-values of the GWB stems, the respective p-values of the two-way ANOVAs were <0.0001 and 0.00058, which means that the effect of both Harvest and Genotype variables on the K-values were significant. PP stems harvested in June or in October performed better than Energo plants. Considering *μmax* values, only the harvest time proved to be a significant factor (p-values <0.0001 and $= 0.0122$), here no difference was found between EN and PP.

On the other hand, one-way ANOVA was used to compare the leaves and stems of the two genotypes separately within each harvest and separately for the two model parameters to find out which genotype is better in the harvests.

Regarding stems the June harvest showed the overall best fermentation parameters (K = 199.08 and 251.80 mL CH_4/g VS; *μmax* = 11.22 and 10.75 mL CH_4/day/g VS, for EN and PP, respectively), the October harvest's values were considerably lower, with the August samples somewhere in between.

In the June harvest PP showed higher K-values than EN (21% increase, p-value = 0.0079), which was in the range of the August samples. These stems were firm, the PP especially, their AD was very efficient: the μmax values were almost in the range of the leaves. In the October samples, PP stems again showed a better performance: 28% higher K- (p-value =

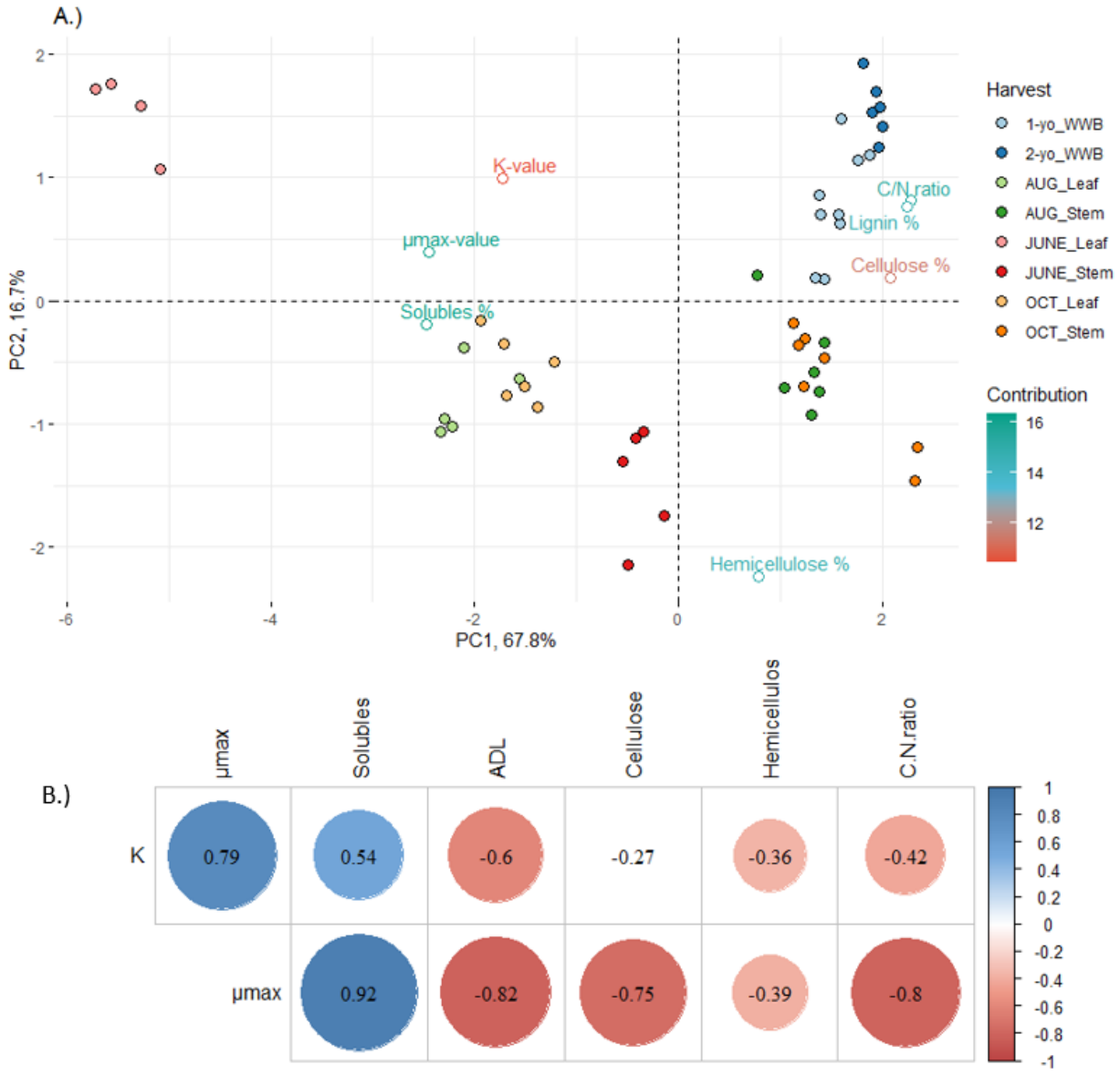
0.0014) and 32% higher *μmax* (p- value = 0.0017) values. In the August samples no significant differences were found between PP and EN stems.

As far as the leaves were concerned, the best values were also obtained from the June harvest, and although in this case there was no difference between EN and PP, the results were remarkably high (mean K = 335.5 mL CH_4/g VS and μmax = 25.4 mL CH_4/day/g VS). These were significantly higher than the other harvest's results (p-values < 0.0001). In the June and Oct samples PP showed a slightly higher K-value (p-value = 0.089) than EN; and there was a slight difference in the case of the leaves' μmax values from October: here EN showed a slight increase in degradation rate (12%). These were the only differences found within harvests, generally only minor variations were observed in these aspects.

Overall, in accordance what was previously hypothesized, it can be argued that the leaf biomass degrades rapidly and yields more methane than either green or woody stems. Moreover, I found that the plant genotype has a significant impact on both fermentation parameters, although to a different extent that depends on the harvesting time.

In order to uncover the reasons, i.e., the molecular backgrounds for the different fermentation efficiencies of the biomass variants, I also characterized their substrate parameters (i.e. fiber and chemical composition). The resulting values are shown in Table 1. The differences between the two genotypes were examined with a (one-way) ANOVA for each parameter independently and the results were subjected to a Tukey's post-hoc test. I also calculated the Pearson correlation values (*Pearson's r*) between the fermentation parameters and the substrate components to find out possible causes for the differences. These correlations are presented in Figure 1B. According to these estimations, both fermentation parameters correlated with the "soluble" content in a positive way: strongly in the case of μmax (r = 0.92, p-value = 0.0002) and moderately in the case of K (r = 0.54, p-value = 0.0006). As this component contains the plant cell's cytoplasm contents, extra- and intracellular oligosaccharides, and proteins as well, the hydrolyzing bacteria can carry out the anaerobic degradation of the solubles component with the

smallest difficulty. In the earlier phases of the fermentation, the higher Solubles content leads to a more rapid biogas production, while the breakdown of the resilient lignocellulose components generally proceeds relatively slowly. The leaves performed considerably better in the batch tests, likely as a consequence of their high concentration of soluble content and low concentration of lignin.



(Please see Figure explanation on the next page)

Figure 1. A): Principal Component Analysis (PCA) biplot of the samples (individuals, filled circles) based on their measured parameters (variables, unfilled circles). This chart shows the overall similarities and differences between the GWB (leaves and stems) and WWB samples, according to their chemical compositions and fermentation parameters. The samples are colored according to the harvest time and to the part of the plant in case of GWB, as shown in the right side of the figure. The X- and Y-axes represent the first and second principal components, respectively. Contribution means the percentage of the variables to the principal components, i.e. how important a given variable is in determining the principal components. PC1 and PC2 describes 67.8% and 16.7% of the overall variability in the data, thus distances in the X-axis are more expressive than in the Y-axis. **B)**: Pearson correlation (*r*) of the fermentation parameters with the substrate parameters. Filled circles represent 'significant' correlations (*p-value* ≤ 0.05). The filling colors of the circles show the scale of the correlations (scale indicated on the right).

I also found that GWB and WWB differed regarding fiber and C/N ratios as well: GWB showed a lower C/N ratio for leaves (average = 15.6), and for stems (average = 19.7), than in WWB (average = 48.1). The values of GWB are closer to the optimal range (Wirth et al., 2018). C/N ratio negatively correlated with μmax and K in a strong (r = −0.80; p-value = 0.00001) and a moderate (r = -0.42, p- value = 0.028) way, respectively (Figure 1B.). However, most of the substrates had a higher-than-optimal values regarding C/N as the C/N = 20–30/1 range has been found the most appropriate for biogas reactors (Böjti et al., 2017). This can be translated to the following: as the C/N value gets closer to the optimal range, μmax values tend to be higher, especially in the case of leaves.

Herrmann et al. thoroughly measured several chemical parameters of many crop silages, but in their results C/N ratio did not show a strong correlation with BMP (Herrmann et al., 2016). This is likely because all these measured substrates were practically in the optimal range. The authors argued that the most important contributor to the AD was acid detergent lignin (ADL), besides acid detergent fiber (ADF), neutral detergent fiber (NDF) and some silage parameters, e.g. butyric acid and alcohols. In their analyses ADL showed a strong negative correlation with K (−0.73), which, according to the results of

my experiments was somewhat less evident, but still significant ($r = -0.60$, p-value = 0.0006). Lower lignin concentrations were found in leaves (mean = 9.6), and in GWB stems (mean = 18.2) than in WWB (mean = 27.8) (Table 1.). The anaerobic biodegradation of lignin proceeds slowly, if all (Nurk et al., 2016), thus pretreating the substrate to decompose lignin is beneficial for the subsequent biogas production from lignocellulosic biomasses, e.g. wheat straw (Rouches et al., 2016b), corn stover (Schroyen et al., 2014) and willow (Alexandropoulou et al., 2017; Jurado et al., 2013). Reviewed by (Patinvoh et al., 2017; Rouches et al., 2016a; Shah et al., 2015). Henceforth I anticipated that higher lignin content will result in less efficient methane production. Indeed, in the present experiments, the slope of the methane production curve was even more affected by the ADL content than the final yields: a strong negative correlation between the μmax values and ADL content was found ($r = -0.82$). This is also supported by the very weak correlation between hemicellulose and K-values, and the practically missing correlation between cellulose and K-values. Perhaps this is the reason why cellulose showed a negative correlation ($r = -0.75$) with μmax, i.e. it degraded during the AD, albeit slower. Rendering the cellulose structure less crystalline-like, is the aim of many other pretreatment processes (Tsapekos et al., 2018). The negative correlation between μmax and cellulose content supports that its rigid structure hinders its decomposition. Most of the cellulose breaks down eventually during the residence time in the AD reactor, hence there is practically no negative correlation between its amount, and the overall CH_4 yield. Hemicellulose was not correlated with the fermentation parameters, usually this substrate is not the bottleneck of the hydrolysis process.

Principal Component Analysis (PCA) was used to highlight similarities and differences between biomass samples on the basis of their chemical compositions and fermentation parameters. The 'individuals' on the biplot (Figure 1B.) are colored according to the harvesting time. The 1- and 2-years-old woody samples (light and deep blue filled circles, respectively) form a distinct group, which is the farthest from the other leaves and stems grouped in between them.

5.1.2. Leaf to stem ratios and methane production in mixed fermentation

In the second set of experiments, I carried out mixed batch fermentations. In these digesters both the leaves (L) and stems (S) were used, and their ratio was adjusted to the L/S ratios of the original plants. The reason for this type of arrangement was that it better represents the scenario, in which the whole plantation of GWB gets harvested for utilization in an industrial biogas digester. The plants had different leaf-to-stem ratios and we suspected that this would influence the K-value and/or the *μmax* values. Figure 2. shows the results and the correlation between the L/S ratios and the two fermentation parameters. In the case of June samples, L/S ratios were higher than 1, which means that in these harvests, leaves contributed to more than half of the total biomass. These values are advantageous for biogas production (see also Table 1) since leaves are readily digestible substrates.

Harvest	Genotype	L/S (oTS)	*μmax*	K
AUG	EN	0.430	7.43	199.4
AUG	PP	0.447	8.33	210.3
JUNE	EN	1.170	12.63	174.0
JUNE	PP	1.220	14.02	199.5
OCT	EN	0.340	7.85	167.2
OCT	PP	0.363	8.22	147.5

Table 2.) Average L/S (on an oTS basis), *μmax* and K values of the mixed fermentations, according to each harvest and plant genotype. In the mixed fermentations the ratios of leaves and stems were put into the fermenters in the same as they were in the original plants L/S ratios.

This observation was supported by the strong positive correlation of *μmax* with the L/S values (Pearson's rho = 0.84), however K-values hardly showed any correlation (r = 0.20). The results indicated that the fermentation rate was higher in the samples where there was more leaf biomass, but apparently the overall CH_4 yield depended primarily on other factors. It is important to note that industrial biogas plants can make substantial advantage

of the utilization of faster degrading feedstocks because in this case the specific daily biogas yields increase.

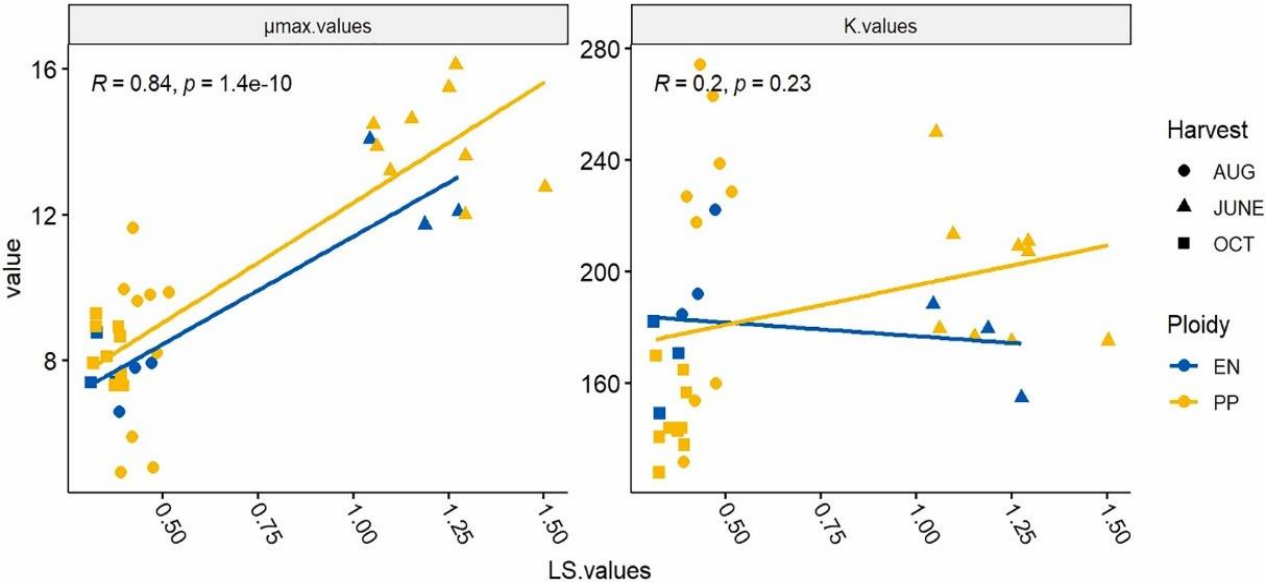


Figure 2.) Correlation of the Leaf/Stem (L/S) ratios (on a VS basis) of EN and PP willow genotypes with the average methane potentials (K) and highest production rates (μmax) derived from the modified Gompertz models of their respective mixed fermentations.

In the mixed fermentations there was no significant difference between the two genotypes, although the tetraploid (PP) biomass harvested in June and August decomposed more efficiently as compared with the diploid plants. Both *μmax* and K values in the October samples were significantly lower than that of the June and August samples. Even though the remarkably high K values of the June samples observed in the monofermentations did not occur here, the *μmax* values were still significantly higher than the other samples' (Figure 2. and Table 2.). It is noteworthy, that Dudits et al (Dudits et al., 2016) found higher L/S ratios in the tetraploid PP green willow plants relative to the diploid ones. The results suggest that the PP GWB may be somewhat better AD substrate than the parent diploid cultivar.

Correlations between the chemical parameters of the substrate and the methane yields (K) and production rates (*μmax*) indicated that low lignin and high soluble contents are the most important factors in high and rapid biomethane yields. These are characteristics of the leaves and the stems from the June harvest and therefore could serve as superior CH_4 sources than the woody tissues. Stems of the PP clones showed higher overall CH_4 yields in the June and October harvests, and faster degradation rates in October samples. Higher leaf-to-stem (L/S) ratios were shown to be positively correlated with *μmax* values, suggesting that PP plants harvested in June and both PP and EN plants from August harvests are the best substrates for biogas fermentation. This can be especially relevant to continuous industrial AD, as these systems take particular advantage of the rate of substrate degradation.

5.1.3. Microbial taxonomic composition analysis of the digesters

The degradation of plant biomass in anaerobic digesters is carried out by a wide variety of microbes, including *Clostridia*, *Bacteroides* and others (Kougias et al., 2018; Wirth et al., 2012). Microbial community analyses can help to better understand these processes and to highlight the most important taxa in the degradation of a specific biomass. In order to carry out such an analysis, that is to identify differences between the genotypes and harvest groups, the microbial compositions of the co-fermentations were examined via 16S rRNA gene-based metagenomics. This was carried out by PCR-amplifying and subsequent sequencing of the V3-V4 hypervariable region of the prokaryote 16S rRNA marker gene from a sample of the total community DNA. The resulting reads were subjected to DADA2's ASV-based workflow, which was shown to have benefits over OTU-based methods (Callahan et al., 2017). These advantages urge ASV methods to be used in biogas related community analyses in future studies – to this we aim to contribute to here as well.

From sequencing results, the read counts of the ASVs in the samples were *rlog*-normalized and a PCA analysis was carried out (see Materials and Methods chapter for details). The results of the PCA demonstrated a clear pattern in the taxonomic composition of the samples (Figure 3.): the microbial communities from each harvest had visibly clustered

together. Although the two PC axes combined represent 55% of the variance of the samples only, no obvious difference or trend between the two plant genotypes could be seen. The similarity of substrate characteristics in each harvest may explain this observation. The clustering of microbial community compositions from June to October was similar, the trend reflected the substrate and fermentation parameters, i.e. that the clustering occurred primarily on harvesting time and the August samples showed a transition between the June and the October samples (Figure 3.).

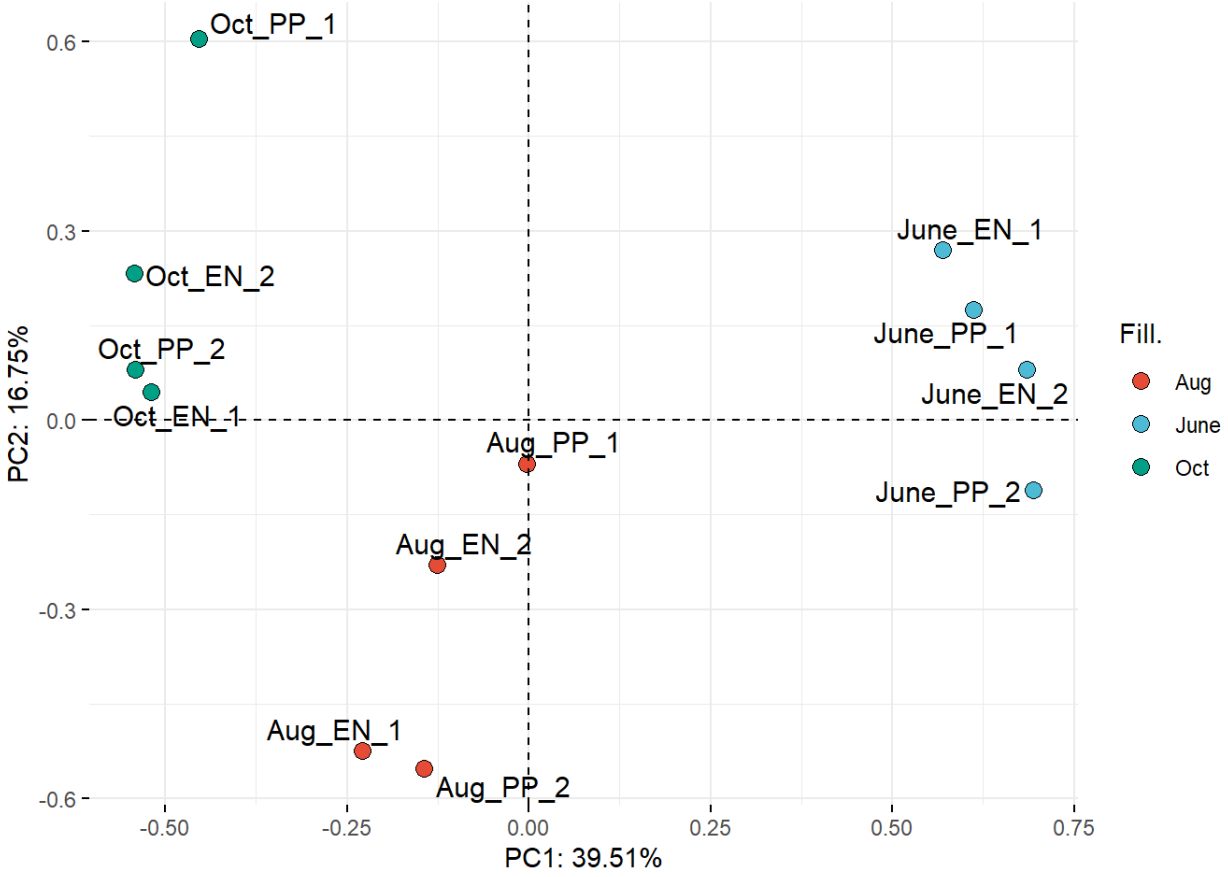


Figure. 3.) Principal Component Analysis presents the established microbial communities at the end of the experiments, according to various harvests (filling colors) and genotypes (labels). The filled circles represent the individual samples, with replicate reactors labeled with numbers.

I investigated the microbial compositions at a higher taxonomic level and found that the order *Clostridiales* was predominant in the samples: their average relative abundance was

80.4% ± 1.0% in the case of August; 80.1% ± 0.3% in October; and 80.8% ± 0.6% in May harvests.

Although this trend was described in other studies as well (Güllert et al., 2016), the practical absence of *Bacteroides* in my samples is unusual. One likely explanation for this observation is that since in this work I carried out an end-point microbial community analysis, the scarcity of metabolic resources and the decreasing pH at the end of the batch fermentation process was better tolerated by *Clostridiales* than *Bacteroides*. *Bacteroides* were present in the initial inoculum but they were gradually outcompeted by *Clostridia*, in a way that is not dissimilar to what happens when temperature is increased in the fermenters (Vanwonterghem et al., 2015).

Two of the most abundant ASVs could be considered as major hydrolytic players in the consortium: *Clostridium butyricum* (ASV_8) and *Herbinix luporum* (ASV_5). The mean relative abundance of the former (ASV_8) was 3.08% ± 0.15%, this population probably contributed to the deconstruction of hemicellulose. It was shown (Jia et al., 2018) that *C. butyricum* expressed numerous key enzymes involved in pentose phosphate pathway, which converts xylose to hexose to direct it to glycolysis. The latter (ASV_5) is a cellulose-degrading microbe from the *Lachnospiraceae* family. In my experimental setup, *H. luporum* [first isolated from a thermophilic biogas reactor (Koeck et al., 2016)] showed a relatively high mean abundance of 4.16% ± 0.37%. Other predominant ASVs included *ASV_1_Unknown_Clostridium_sensu_stricto_1*, and *ASV_2_Unknown_Clostridium_sensu_stricto_1*, in average they represented 12.89% ± 0.27% and 9.85% ± 0.28% of the total community, respectively, all from the genus *Clostridiaceae*. Jia et al. has observed that transcriptionally very active microbes (Jia et al., 2018) could contribute to as much as 60–70% of the total community abundance. It is reasonable to assume that a similar phenomenon occurred in the present system as well.

As the PCA results on the overall taxonomic composition showed that samples have clustered together in accordance with the harvesting time and effects that could be contributed to genotype were insignificant, I carried out a differential abundance analysis

with the DESEq2 workflow on the ASVs as a factor or harvest groups (Figure 4.). This was to reduce the number of comparisons, which increases the statistical power of the tests. These results showed a similar pattern that was observed in the PCA results, that is the highest number of statistically significantly different populations (DPs) were in the June-October comparison (23), while the August-October comparison yielded the lowest (5), with the August-June comparison in between (13).

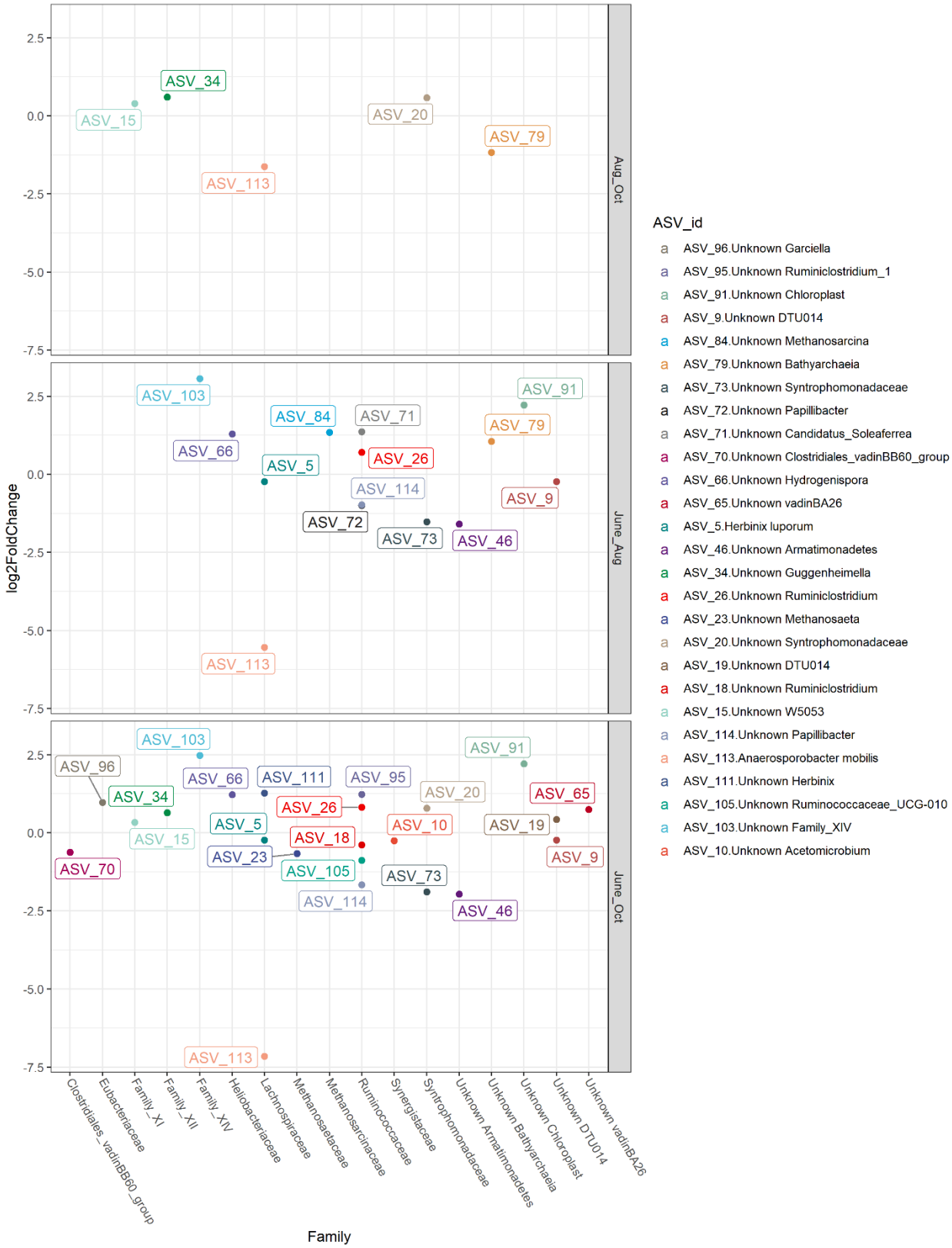

log2FoldChange
Aug_Oct
June_Aug
June_Oct
Family
Clostridiales_vadinBB60_group
Eubacteriaceae
Family_XI
Family_XII
Family_XIV
Heliobacteriaceae
Lachnospiraceae
Methanosaetaceae
Methanosarcinaceae
Ruminococcaceae
Synergistaceae
Syntrophomonadaceae
Unknown Armatimonadetes
Unknown Bathyarchaeia
Unknown Chloroplast
Unknown DTU014
Unknown vadinBA26
ASV_id
ASV_96.Unknown Garciella
ASV_95.Unknown Ruminiclostridium_1
ASV_91.Unknown Chloroplast
ASV_9.Unknown DTU014
ASV_84.Unknown Methanosarcina
ASV_79.Unknown Bathyarchaeia
ASV_73.Unknown Syntrophomonadaceae
ASV_72.Unknown Papillibacter
ASV_71.Unknown Candidatus_Soleaferrea
ASV_70.Unknown Clostridiales_vadinBB60_group
ASV_66.Unknown Hydrogenispora
ASV_65.Unknown vadinBA26
ASV_5.Herbinix luporum
ASV_46.Unknown Armatimonadetes
ASV_34.Unknown Guggenheimella
ASV_26.Unknown Ruminiclostridium
ASV_23.Unknown Methanosaeta
ASV_20.Unknown Syntrophomonadaceae
ASV_19.Unknown DTU014
ASV_18.Unknown Ruminiclostridium
ASV_15.Unknown W5053
ASV_114.Unknown Papillibacter
ASV_113.Anaerosporobacter mobilis
ASV_111.Unknown Herbinix
ASV_105.Unknown Ruminococcaceae_UCG-010
ASV_103.Unknown Family_XIV
ASV_10.Unknown Acetomicrobium

Figure 4.) The *log2* Fold Change (Y-axis) of significantly different ASVs between each harvest. Positive values indicate the that the given ASV was signifantly more abundant in the first sample of the comparison (e.g. June in the June_Aug comparison). The ASVs are grouped by taxonomic Family (X-axis). The three panels show the three comparisons between the harvests.

5.1.4. The advantages of using green willow biomass in comparison to other biogas feedstocks

Biomass yields per unit area (e.g. hectare) must be taken into consideration for an appropriate evaluation of the bioenergy potential of a novel biomass as a feedstock for large scale AD (Gissén et al., 2014). In order to estimate the biomass yields of GWB, I carried out a small-scale field experiment. It should be noted that this approximation may not reflect precisely yields for the large scale plantations, but nevertheless the trend should be informative. To determine the optimal harvest time, I estimated the biomass yields at each harvest time point and compared to that of WWB and maize silage (Figure. 5.A). I used the EN genotype for this comparison, as it is a commercial and publicly available willow cultivar. I harvested the willow biomass for biomass yields estimation in four occasions: May, June, August, and October; and for each harvest, 10 different plants were analyzed. The willow biomass grew steadily during Spring and Summer before starting to decrease mainly due to loss of leaves during the second half of September and in early October. Since after the harvest in August the biomass was apparently still accumulating, I carried out a prediction for an early September harvest as well, by fitting a growth model on the measured data. I compared GWB biomass yields to four yield data for maize (Fuksa et al., 2020; Gissén et al., 2014; Jankowski et al., 2020; Wannasek et al., 2019) and one for WWB (Stolarski et al., 2017) from the relevant scientific literature. The measured and the estimated biomass yields of GWB were in the similar range or somewhat higher than that of WWB or maize (Figure 5.B), supporting the correctness of the estimations.

As mentioned before, the extractable energy potential is of main importance when assessing the possible utilization of a novel biomass and/or biogas substrate for industrial

applications. Thus, I estimated the per hectare biomass yield (t VS / ha) from the actual (measured) yield values of the harvested plants, and then multiplied this with the substrate specific methane yields (mL CH_4 / g VS) and the energy content of methane to get the per hectare energy potential. These estimates were then compared to the energy potential of other biomasses (Figure 5.A.).

According to my estimates, GWB harvested in August and September can produce biomass yields (22.8 and 31.5 t/ha, respectively) that are significantly higher (on a VS basis) than either WWB or maize; and biomass yields of GWB harvested in June and October are comparable to these (11.7 t/ha and 14.1 t/ha, respectively).

Maize silage evidently has a higher VS-specific methane yield (mL CH_4 / g oTS) than GWB (or WWB), thus regarding the per hectare energy potential there were noticeably less differences between the feedstocks (Figure 5.B.). In the case of maize silage, the energy potential per plant was estimated by dividing the per hectare energy potential taken from literature data with the used planting density. In the case of GWB however, it was the other way around: the measured values from the plants were scaled up with the planting density to get Ep/ha (per hectare energy potential). This is the reason why there is an apparent linear connection between the per plant and per hectare energy potential in GWB but not in maize.

This comparison revealed that GWB can be a valuable substrate for biogas generation. The Ep/ha values in the case of August harvest and September estimation were 176.27 and 227.65 GJ/ha/year, respectively, which are in the range of the values of maize silage and are higher than that of WWB when it is used in anaerobic digesters and similar to WWB when it is used for burning.

It should be further emphasized that compared to WWB, GWB performs better from an industrial-scale point of view not only due to its higher Ep, but also because its higher μmax values. This is because faster CH_4 production is more advantageous in CSTR reactors.

The results showed that GWB can reach the average Ep values for maize silage biomethanation, although the specific yields depend on planting density, which can be as

dense as 90,000 GWB stem/ha or 130,000/ha (Fuksa et al., 2020), and possibly other factors as well. If agro-technologies would be optimized for higher plant densities, it is very likely that the energy potentials of GWB plantations could be elevated more and thus even surpass that of a maize silage. All this for a substantially lowered production cost. And while to ascertain the exact amount is out of the scope of this work, according to a personal conversation with the CEO of the Kaposszekcső Biogas plant, whom we collaborated with and who grows both willow and maize, willow is about 3-4 times cheaper to produce. Nevertheless, additional validations of the biomass yields from larger-scale and multi-year field experiments are necessary to gain higher confidence in the biomass yields its production costs, and to evaluate the possible attenuation of annual cutbacks of the plants.

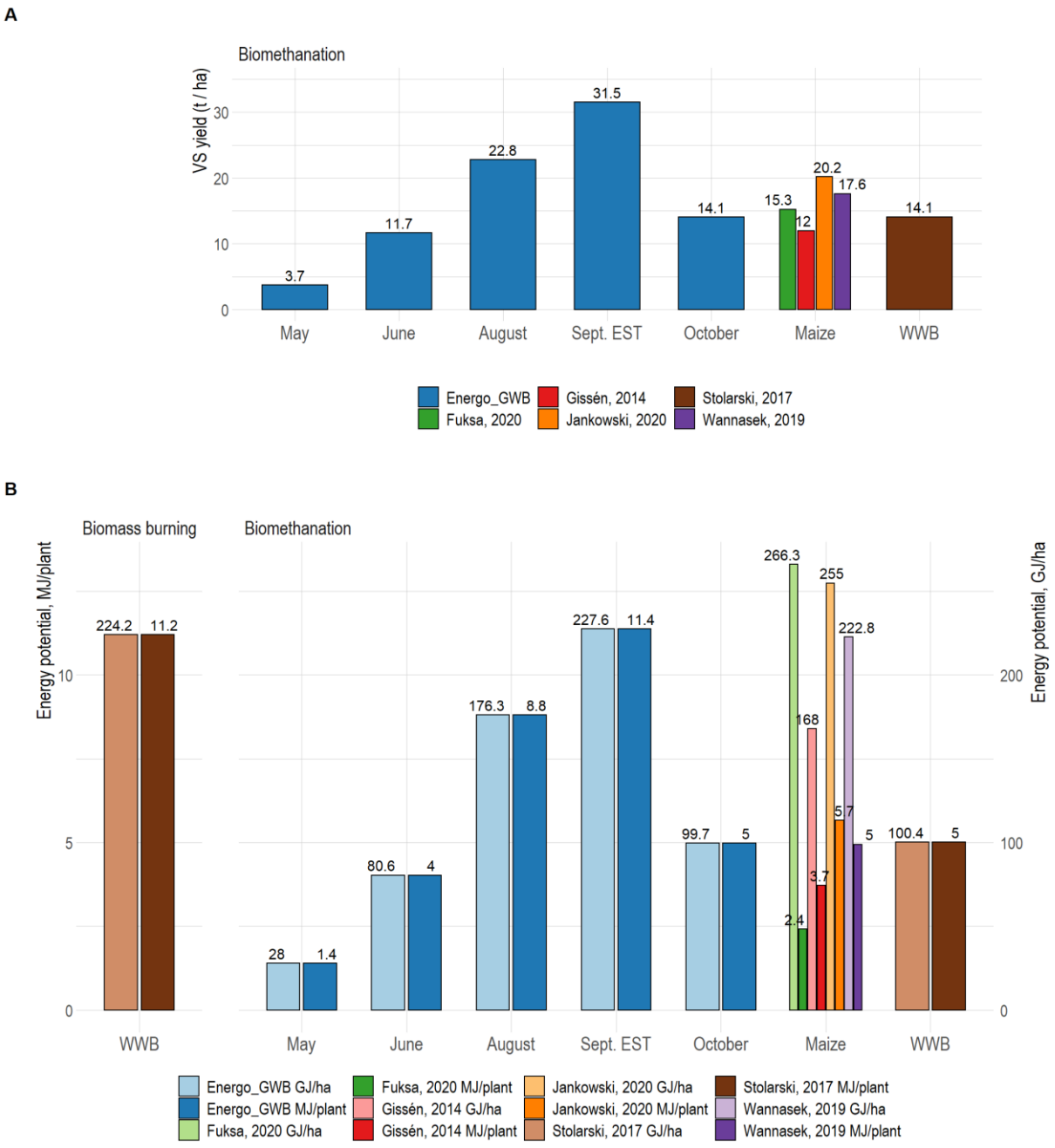


Figure 5.

A): Extractable energy potential (Ep) of various substrates via biomethanation based on plantation area (GJ/ha, lighter colors) and individual plant (MJ/plant, darker colors). The energy potential of the plants and the plantation in each harvest was calculated by multiplying the VS yields with the

substrate specific CH4 yields, as described in Section 2.1. The values for maize silage and WWB are from the respective references.

B): Green willow biomass (GWB) yields (t/ha) at various harvesting times. The VS yield/ha was calculated by scaling up the mean of the individual VS yield of the Energo plants, with the used planting density (20,000 plant/ha). To estimate the biomass yield in early September (Sept. EST), a Gompertz growth model was fitted to the measured data and the values predicted by the model was used. Biomass yield values for maize are from: (Fuksa et al., 2020; Gissén et al., 2014; Jankowski et al., 2020; Wannasek et al., 2019) and for woody willow biomass (WWB) from (Stolarski et al., 2017).

5.2. H_2-induced stress response analyzed by combined metagenomics and metatranscriptomics

My aim was to investigate which genes and pathways are activated firstly in the anaerobic consortium – and mainly in the methanogenic Archaea – in response to H_2 addition. In the experimental setup I used, an excess amount of external H_2 was injected to a stable biogas producing anaerobic sludge that was fed with alpha-cellulose. I analyzed the effects of the sudden H_2 burst on the microbial community at a metagenomic and metatranscriptomic level. Figure 6. shows the general experimental arrangement and the main analysis steps involved.

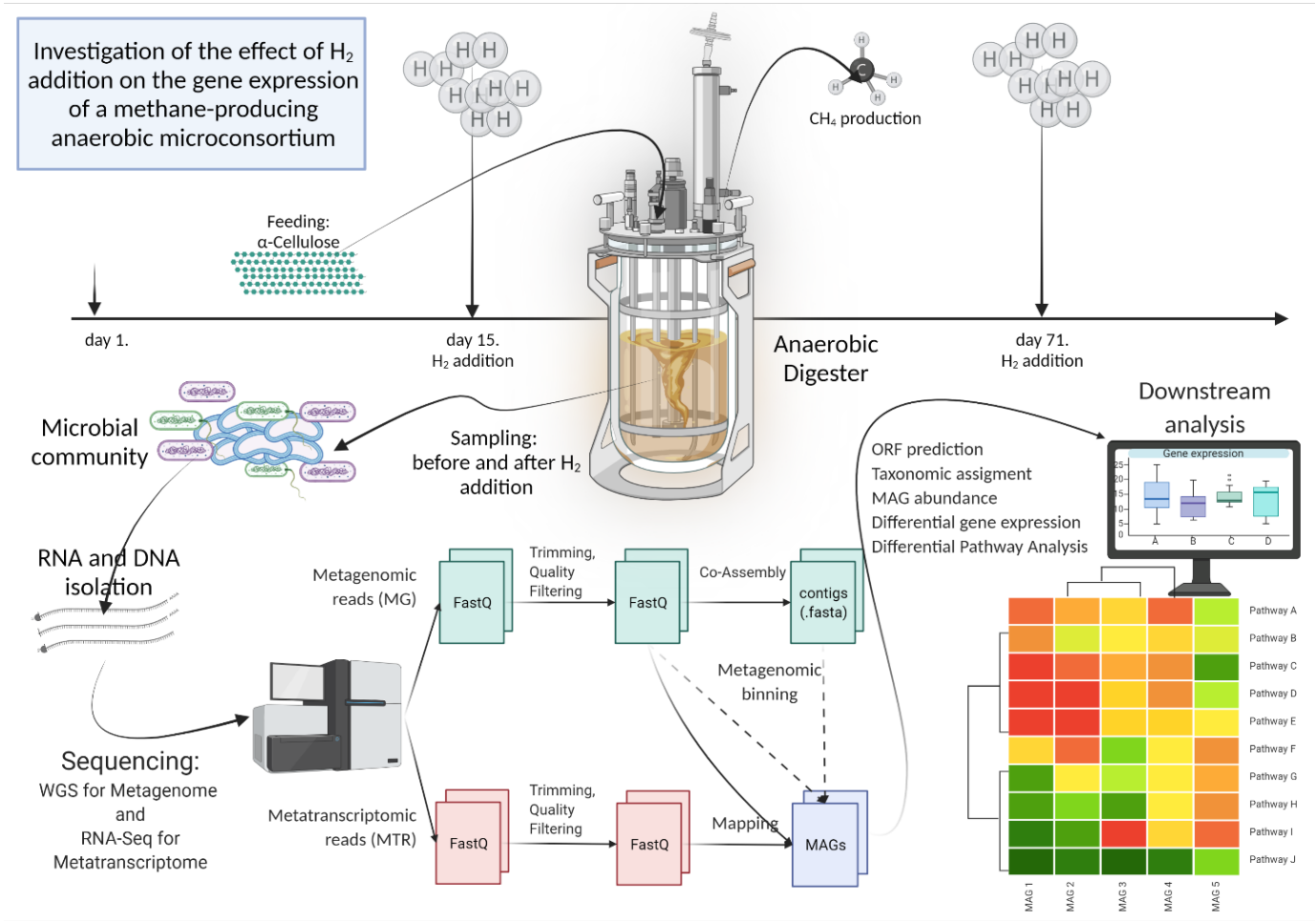


Figure 6.) General experimental arrangement and the main analysis steps. A 5 L CSTR reactor producing stable biogas values was injected with H2 for 10 minutes. Samples were taken from it, immediately before this and 2 hours later. Total community DNA and RNA was isolated from the samples and were subsequently sequenced. Metagenome-assembled genomes were generated and analyzed for function, taxonomy and abundance. The detailed bioinformatic pipeline is described in the Materials and Methods section.

I started the fermentation with a 5-days long start-up phase, during which no substrate was added. Subsequently, a 15-days-long adaptation phase followed, during which alpha-cellulose was added in a quantity of 1 g/L/day. Achieving a balanced operation was desired, thus a relatively low substrate loading rate was used. To monitor the fermentation process, I assessed VOA/TIC and NH_{4+} levels, measured biogas productions and methane concentrations. During this period the average VOA and TIC values were close to 1.1 g/L and 14 g $CaCO_3$ /L. This established that the process was indeed balanced, as a steady value of VOA/TIC is a reliable indicator of a stable mesophilic fermentation (McGhee, 1968). NH_{4+} level is another important indicator of AD process stability (Alexander Martin, 1985). Theoretically, levels above 3,000 mg NH_{4+} may have a negative effect on the methanogenic archaea, which is the most sensitive group of microbes in the AD process (Yenigün and Demirel, 2013). The NH_{4+} concentration was below 1,000 mg/L during the whole fermentation process. The biogas productivity of the digesters was also stable: on average 650 L_N biogas was produced from alpha-cellulose/VSg/day with an average of 53% CH_4 content, the daily biomethane production varied by < 10%. Thus overall, the reactors showed stable operation during the course of the experiment (Supplementary Figure S1).

The first two DNA and RNA samples for metagenome and metatranscriptome analysis were taken on day 20 from the stabilized reactors. Subsequently the digesters were flushed with H_2 gas from a gas cylinder for 10 min. This was to imitate a power-to-gas system wherein H_2 resulting from water electrolysis is added as a means of biological biogas upgrading. Two hours later I carried out a second sampling. I repeated the protocol after

2 months of reactor operation (on day 71) and treated the respective samples as biological replicates.

The reactors responded with a sudden increase in daily CH_4 evolution by 20–25% at both time points, which lasted for 1–2 days (Supplementary Figure S1). Afterwards the reactors returned to their previous biomethane production levels (~53%). It is worth noting that the bioreactor system responded almost in exactly the same manner to the H_2 spike 2 months apart, which indicates the robustness, reproducibility and quick response time by the microbial community. Assuming H_2 saturation of the liquid phase by the 10 min long H_2 bubbling, I estimated that more than 95% of the injected H_2 was converted into CH_4 within 16–24 h, although the amount of available dissolved H_2 decreased rapidly during the second half of the H_2 consumption phase. This was in line with the observations of Szuhaj et al. (Szuhaj et al., 2016), who found that in fed-batch systems the injected H_2 was completely consumed in 16–24 h despite much lower saturation. Except for the peak CH_4 production right after H_2 injection, the added H_2 apparently did not alter markedly the cumulative biomethane production.

5.2.1. General overview of the genome-resolved metatranscriptomics pipeline

As for the reasons explained in the *Introduction* section, combining MTR with MG yields a real-time snapshot of the functioning of the microbial system, which can provide more insight into it than when either is carried out separately. Since my main interest was pointed towards the reaction of individual microbial populations (aka MAGs) to the H_2 addition, and since none of the currently available bioinformatics tools offer a complete workflow for that end, I decided to create and use a custom bioinformatic pipeline. The resulting MTR and MG reads were thus analyzed with following aims:

1. Generate metagenome-assembled genomes (MAGs) from the MG reads via binning the contigs resulting from a co-assembly of reads.
2. Carry out ORF prediction and functional annotation on the MAGs, involving gene function and pathway analysis (based on the KEGG).

3. Taxonomically annotate the MAGs and examine their phylogenomic relationships.
4. Map both MG and MTR reads to the MAGs to estimate gene copy number (MG) or gene expression (MTR).
5. Find differentially expressed MAGs, ORFs and pathways between the N_2 and H_2 MTR samples.
6. Analyze and visualize the results.

This workflow relies heavily on the following tools, methods: Anvi'o (Eren et al., 2015), DASTools (Sieber et al., 2018), SqueezeMeta (Tamames and Puente-Sánchez, 2019), Clusterprofiler (Yu et al., 2012) and DESeq2 (Love et al., 2014). The general steps are shown in Figure 6. and the details can be found in the Materials and Methods section.

When investigating the early response of the residing microbial consortium to sudden H_2 burst at the transcriptome level of metagenome-assembled genomes (MAGs), I assumed that the underlying microbial consortium will not change significantly at the DNA-level, due to the short time interval applied between the samplings (2 hours). Increasing the number of metagenomic samples thus enabled an extensive binning procedure – using all 4 MG samples for the co-assembly. Therefore, the H_2 triggered differences in the gene expression levels could be precisely assessed, together with associated alterations in cell physiology. RNA samples taken before the H_2 addition served as a 'control', while the samples taken after it served as the 'case' conditions for the metatranscriptomic analyses (differential gene expression, differential pathway abundance, etc.), hereafter named N2-MTR and H2-MTR.

5.2.2. Metagenomic binning results and taxonomical annotation

The extensive binning procedure, employing three metagenomic binning programs (Alneberg et al., 2014; Kang et al., 2015; Wu et al., 2016), whose results were refined by a bin quality-improving program (Sieber et al., 2018), yielded a final number of 84 bins (Figure 7.). Each of the initial binning procedures generated more than that, thus this final set will be referred to as MAGs henceforth, to distinguish the two sets. CheckM (Parks et al., 2015) was used to estimate the completeness and redundancy of the bins based on the

prevalence of single-copy marker genes (SCG). Based on these values, out of the 84 MAGs, 16 were high, 49 were medium and 19 were low quality, according to the MIMAG initiative (Bowers et al., 2017). The phylogenetic relationships of the MAGs were assessed with phylophlan3 (Asnicar et al., 2020), and the results indicated that 73 bins harbored enough SCGs for the phylogenetic tree building (dendrogram in Figure 7.)—the phylogenetic relationship of the remaining 11 bins could not be determined, either because of the low quality of the metagenomes or because they did not contain enough known SCGs. The taxonomy of the MAGs was determined using the GTDB toolkit (Chaumeil et al., 2020).

According to the taxonomic analyses, 7 MAGs could be assigned as Archaea, 61 as Bacteria while 16 MAGs remained unclassified (details are presented in Supplementary Table 1.). Archaea represented about 10% of the microbiome, all of them coming from the phylum *Euryarcheaota*. Within the domain Bacteria, most MAGs (34) were associated with the phylum *Firmicutes*. The dominance of *Firmicutes* in biogas reactors is in accordance with previous studies (Campanaro et al., 2020; Wirth et al., 2012). This can be attributed to their diverse polysaccharide and oligosaccharide degradation capacities in, which is the first step in the AD of complex organic substrates (Güllert et al., 2016).

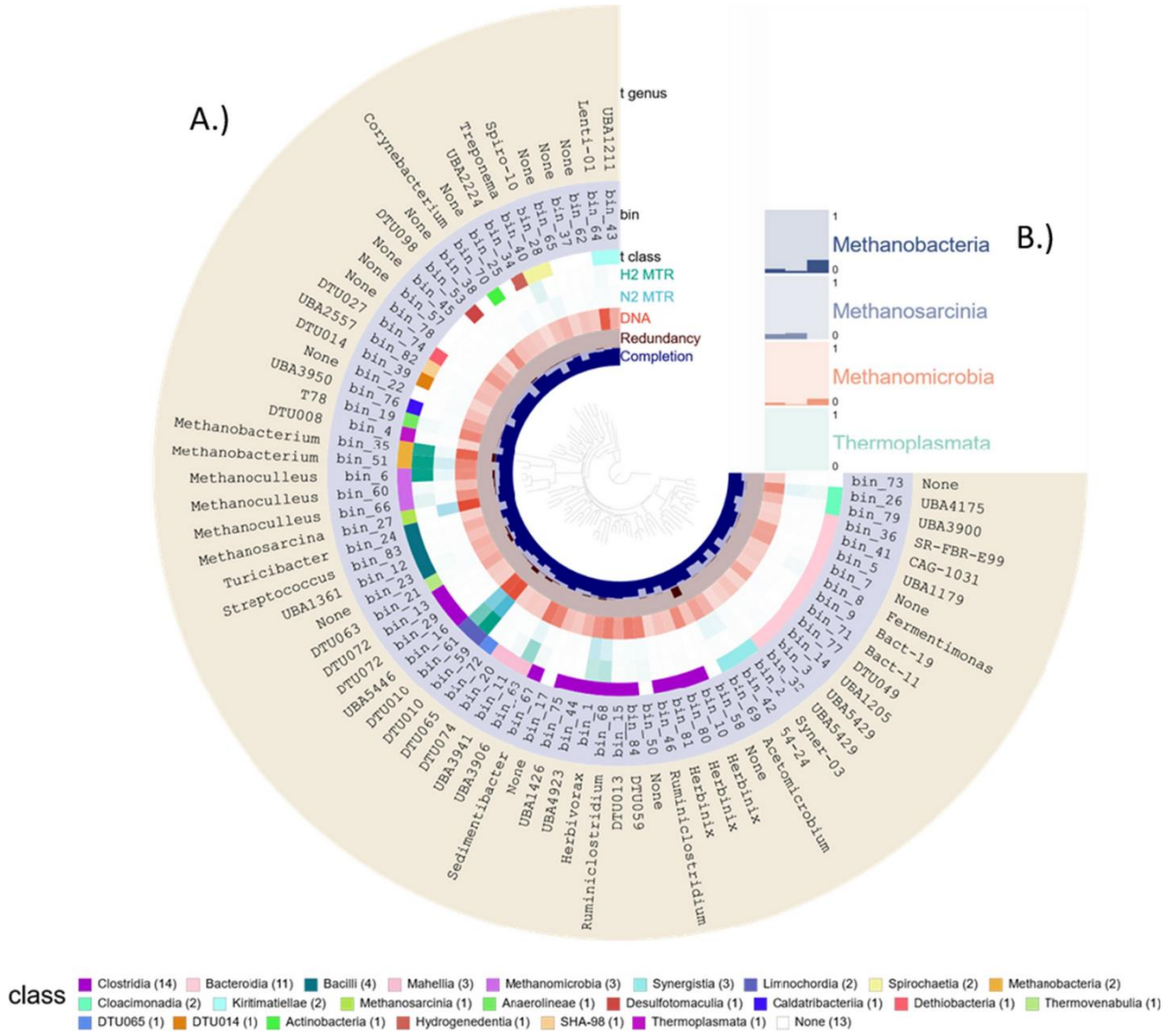


Figure 7. A. Anvi'o plot of binning results (from innermost to outermost): 1.) dendrogram of the phylogenetic relationship of MAGs (according to phylophlyan3); 2.) completion and redundancy estimations (according to CheckM); 3.) relative abundance of the MAGs in the three datasets (MG, N2-MTR and H2-MTR); 4.) taxonomic Class assignment for the MAGs (based on GTDB); 5.) MAG (bin) ID; 6.) Genus assignment for the MAGs. The list of Classes at the bottom part indicates the color code and the number of MAGs in the Classes. **B)** depicts the relative abundance of Archaeal Classes (the total abundance of MAGs in the respective Classes).

The second well-represented phylum were *Bacterioidetes* (12 MAGs), all belonging in the order *Bacteroidales*. Most *Bacterioidetes* produce succinic acid, acetic acid, and in some

cases propionic acid, these molecules fuel acetotrophic methanogenesis. In addition, representatives of the phyla *Synergistota*, *Spirochaetes*, *Verrucomicrobia*, *Cloacimonadota*, *Fibrobacterota*, *Caldatribacteria*, and *Chloroflexota* were identified (Supplementary Table S1). At lower taxonomical levels, the MAGs are similar to those described in previous studies (Campanaro et al., 2018; Treu et al., 2016b).

5.2.3. Comparison of the MAGs with the Biogas Microbiome (BGMB) database

A comparison with the MAGs from the *biogas microbiome* collection (Campanaro et al., 2020) showed that many MAGs in my dataset are unique and cannot be found in other fermenters – or to be more exact, their genomes could not be assembled from that collection of metagenomes. The comparison was based on the average amino acid identity (AAI) of the protein coding sequences and was carried out on the MIGA platform (Rodriguez-R et al., 2018). The results are shown in Supplementary Table S2. Based on this similarity measure, 56 MAGs showed more than 95% identity (Figure 8., black lines), while the other 28 genomes had only 80% or lower identity to any of the 1635 *biogas microbiome* genomes, out of these 14 had no corresponding genome in the database. Thus, there is a noticeable gap between the distribution of maximum similarities: there was not a single MAG between 80% and 95%. This suggests that the species delineation occurs at 95% AAI, while the genus boundary is somewhere between 70-80% AAI. These values are in line with empirical data proposed earlier (Konstantinidis and Tiedje, 2005; Rodriguez-R and Konstantinidis, 2014). The hits between 70-80% probably come from the same genus as the best hit. Below 70%, the corresponding MAGs are probably match only at Order or Class level only. This means that 5 MAGs come from genera that are not included in the biogas microbiome dataset yet, and possible 9 further ones are members of novel Orders and/or Classes. The remaining MAGs contained medium or low-quality metagenomes, which could be the main reason for not finding matches. With about 40% of redundancy, bin_5 (from order *Bacteroidales*) seems to be a mixed genome of possibly two or maybe even three populations. The high identity of the two PGs probably hindered a correct binning procedure. The same might be true for the other highly redundant MAGs as well.

However, these MAGs that have no correspondent in the biogas microbiome database are not always just side-players, rather several of them are abundant, main-players of the community. This includes two high quality MAGs: bin_1—*Herbivorax saccincola* and bin_35—*Methanobacterium* sp. These MAGs were apparently not represented among the more than 1600 PGs identified in Biogas Microbiome. What is more interesting is that these two MAGs are major contributors in my microbiomes, with relative abundances ranging from 5% to 15% and the case of bin_1, and up to ~22% in the case of bin_35, H2-MTR samples. Bin_59, annotated to be a member of Class *Limnochordia*, was one of the most abundant members of the community, reached 15% relative abundance in the N2-MTR sample. The low quality of these MAGs not always seem to be the consequence of low genome coverage, but rather that two highly similar genomes co-exist, and their populations are combined into one during binning. This is supported by the high variance in their abundances: bin_35 in H2-MTR, and bin_59 in N2-MTR.

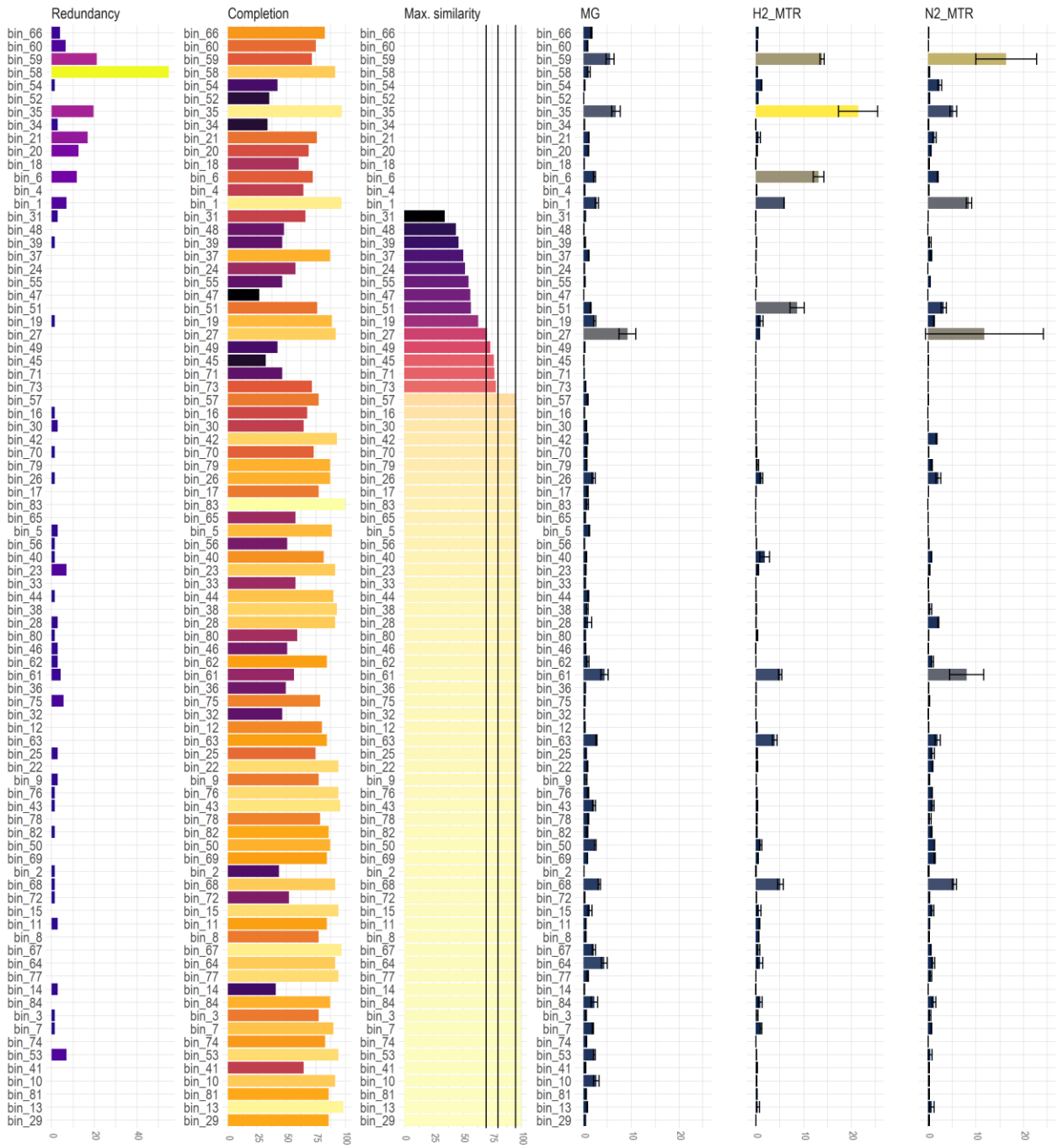


Figure 8.) MAG quality statistics and relative abundance. From left to right: Panel 1.: Redundancy; Completion; Maximum identity: the AAI with the best hit from the *biogas microbiome* collection of 1635 MAGs; MG: mean relative abundance of the respective MAG in the four metagenome samples, H2- and N2-MTR: mean abundance of the respective MAGs in the corresponding metatranscriptomics samples. AAI of 70%, 80% and 95% are indicated by black vertical lines.

5.2.4. Abundance of the MAGs

The metagenomic data clearly indicated that the MAG compositions were very similar in all four MG samples. I carried out a Principial Component Analysis (PCA) on the *rlog*-transformed MAG abundance matrix (please refer to Materials and Methods section) to evaluate the coarse-scale differences between the samples (Figure 9.). The two main observations that could be drawn as correct were as follows *i.)* the prediction that two hours was enough for the MTR samples to diverge significantly, and *ii.)* meanwhile the MG samples remained effectively unchanged. The results confirmed that the microbial community did not change perceptibly within 2 hours even after a substantial change in environmental conditions (H_2-injection) and the experiment was highly reproducible after a 2-month time interval.

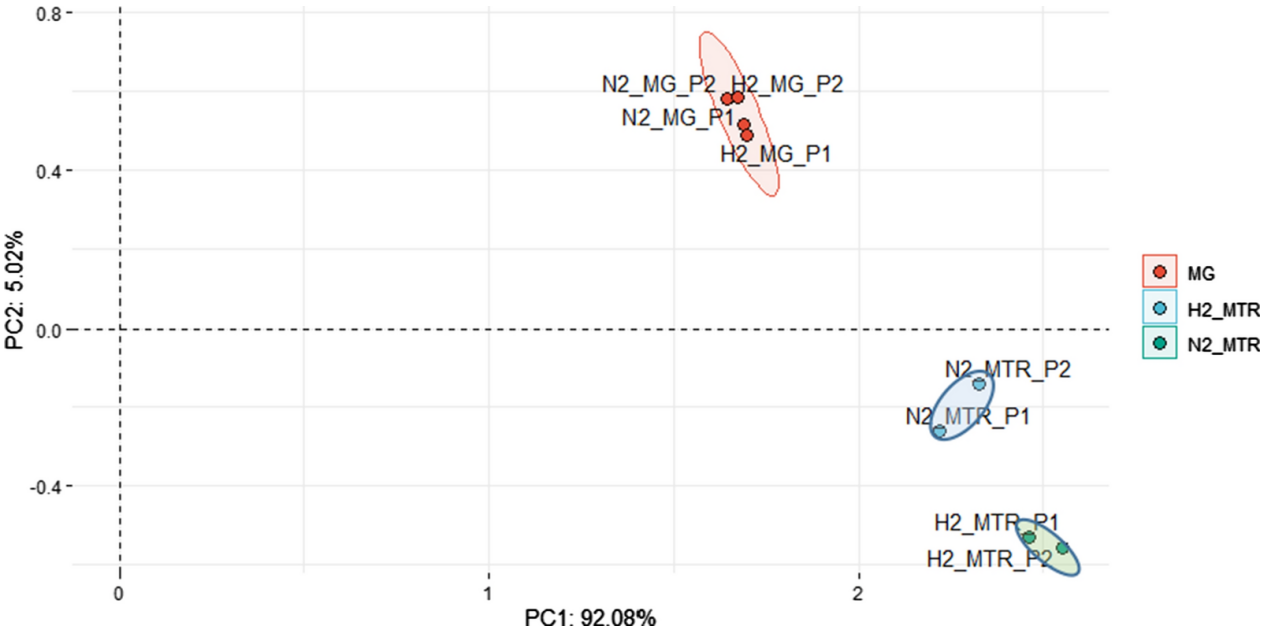


Figure 9.) PCA biplot of the rlog-transformed (regularized-logarithm transformation) MAG abundance in each sample. MAG abundance was calculated as the sum of the reads assigned to the ORFs of the respective MAG, as is the total gene copy number in MG and total gene expression in MTR samples.

I evaluated the MAG abundance in the MG samples at Class-level, which showed that not only the number of assembled MAGs was the highest in *Clostridia*, but their combined abundance as well (Figure 9.). Their combined relative abundance was 19.3% (an average of the four MG samples, sd=0.8%). In this regard, the microbial community thriving in our biogas reactors, was similar to the MAG distribution described by Zhu et al, where in a similar arrangement, the fermenters were fed with Avicel cellulose (in the first feeding phase, later it was replaced with increasingly simple substrates) (Zhu et al., 2019). Regarding the other taxa, the similarities did not hold, as for example *Bacteroides* and *Verrucomicrobiota* PGs were essentially missing in that microflora (or were represented sub-threshold abundances), while in the current dataset, they were abundant (Figure 10.).

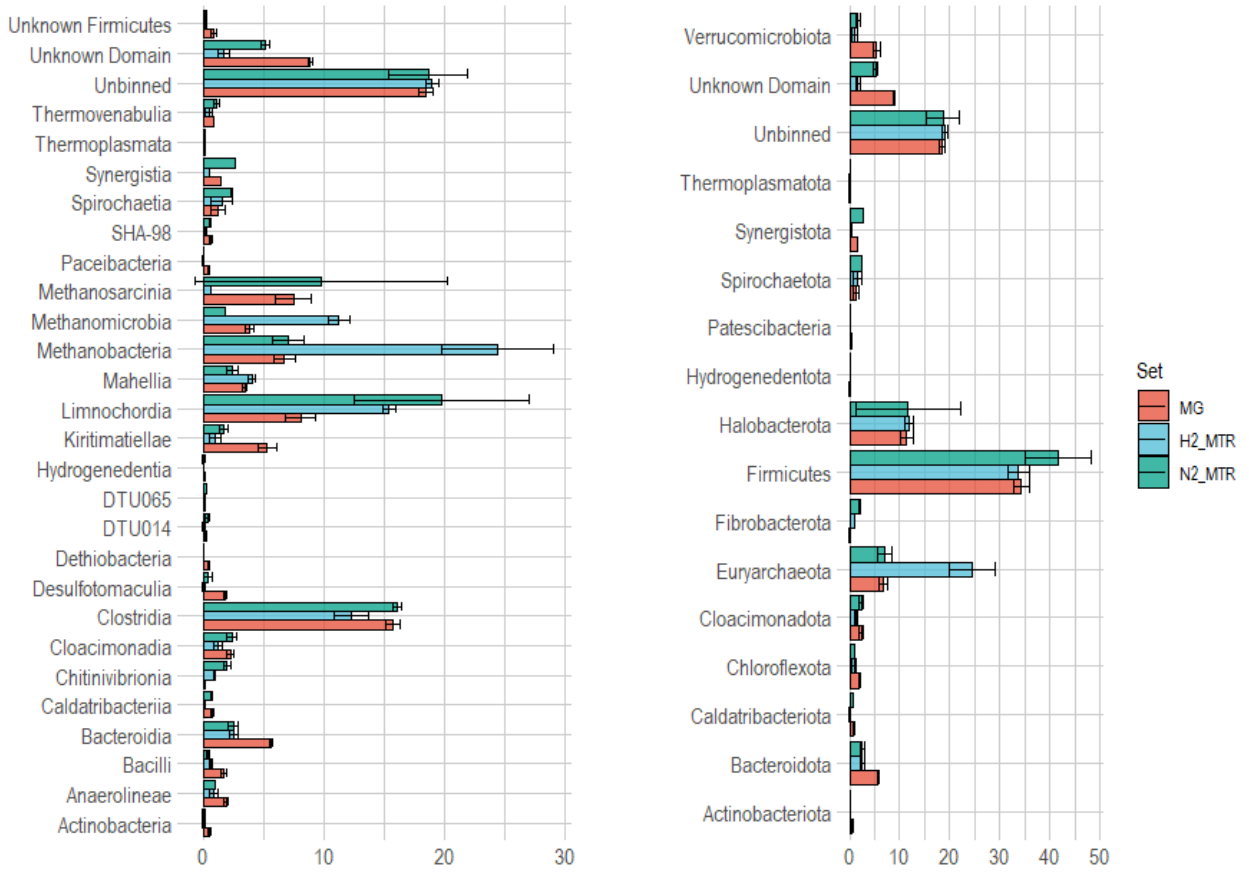


Figure 10.) Taxon abundance in the fermentors, according to the sample types (MG, N2-MTR and H2-MTR). Left panel shows the combined abundance of MAGs at *Class* level, while right panel shows it at *Phylum* level. Abundance was calculated based on the number of reads assigned to the ORFs of the MAGs.

The metagenomics data showed that at DNA-level the community didn't change perceptibly. For example, the overall Archaea gene abundance, i.e., sum of read counts, was 18.49 ± 2.04% in N2-MG or H2-MG samples (each DNA-based). In contrast, the mRNA-based metatranscriptome analysis showed striking changes in the transcriptome-based community composition when H_2 was offered to the reactors' microbial community. The N2-MTR samples (RNA-based, before H2 addition) showed a similar total Archaea abundance to that of the MG samples: 18.99 ± 11.64%, but this was elevated to 36.53 ± 3.74% in the case of H2-MTR samples (RNA-based, after H_2 addition). This demonstrates a rapid response to the appearance of excess H_2. Regarding Archaea, perhaps the most noticeable difference was that in the MG samples (DNA-based data) and N2 samples the Euryarchaeota abundance was very similar, but it elevated more than 3-fold (from about 6.5-7% in MG and N2-MTR samples to 24.4% in H2-MTR samples. This can be attributed mainly to the elevated gene expression of *Methanobacteria* MAGs, and the simultaneous decrease of *Methanosarcina* MAGs.

To infer a more precise comparison, I estimated the log2 fold change (*log2FC*) for each MAG (a measure of differential abundance between the control a test samples) via DESEq2 (Love et al., 2014). This showed that the elevation of the total number of transcribed Archaea genes (H2-MTR samples) was mainly attributed to representatives of the genus *Methanobacterium* (bins 35 and 51), which increased from 4.33 to 17.39% (*log2FC* = 2.84) of all MAGs' abundance. *Methanobacteria* are hydrogenotrophic methanogens. The second major contributor to Archaea transcripts was the order *Methanomicrobiales*, from 2.69 to 7.03% (*log2FC* = 2.16). The genera *Methanoculleus* and *Methanosarcinia* both belong in this order. The three MAGs of the genus *Methanoculleus* showed elevated overall abundance, the *log2FC* values of bin_6, bin_60

and bin_66 were 3.64, 2.37 and 2.18, respectively. The increase upon H_2 exposure was the most apparent in the case of *bin_6. Methanoculleus bourgensis*, whose proportion increased from 1.65% to 10.66%. Remarkably, the genus *Methanosarcina* effectively ceased to express genes upon H_2 dispensation. *Methanosarcina* are known to possess genes coding for all three methanogenic pathways, i.e., hydrogenotrophic, acetotrophic and methylotrophic methanogenesis (Ács et al., 2015; Wirth et al., 2012). Members of the genus *Methanoculleus* are solely hydrogenotrophic methanogens. H_2 exposure apparently turns on the activity of the hydrogenotrophic methanogenesis in both *Methanoculleus* and *Methanobacterium* but turns off the hydrogenotrophic pathway in *Methanosarcina*.

5.2.5. Metatranscriptomic pathways analysis

A community-level pathway enrichment analysis (PWE) was performed to examine the overall metatranscriptomic changes that occurred as a result of the H_2 addition. The input values for this analysis were the log2 fold change (*log2FC*) for each annotated ORF of every MAG, which were also assessed using the DESeq2 method. The analysis was carried out by the *clusterprofiler* R package based on KEGG Pathways and Modules. Although the contig assembly and subsequent ORF prediction and annotation yielded 219,353 KEGG Orthology (KO) annotated ORFs, less than half of them (98,791 ORFs) were binned in the refined MAGs. Thus, despite the binning efforts, the remaining 120,562 KEGG annotated ORFs remained unbinned. However, omitting these from the downstream evaluation would have distorted the statistical analyses, therefore they were grouped together into an *unbinned* collection and were included in the PWE.

The results indicated that the primarily affected KEGG Pathways due to H_2 injection were *Methane metabolism* and *Ribosome* (Figure 11.). The upregulation of differentially expressed (DE) genes was the highest in these modules and their associated KEGG Modules: 48 and 44, respectively. It is noteworthy that some other carbon metabolism associated pathways were also significantly impacted, such as Glycolysis/Gluconeogenesis and Propanoate metabolism, which suggest that acetogenic and acetate utilizing microbes

were also affected by the specifically altered environment. H_2 is known to inhibit acetogenic microbes (Batstone et al., 2006), thus their response to the H_2-addition was expected. The RNA polymerase pathway also changed significantly; this was due to triggered transcription machinery as a response to the altered environment. The KEGG Module enrichment analysis showed a similar picture, the modules that were affected the most were the different methanogenesis modules. This clearly indicates that the two-hour interval was enough for a significant transcriptome-level alteration of the microbial community.

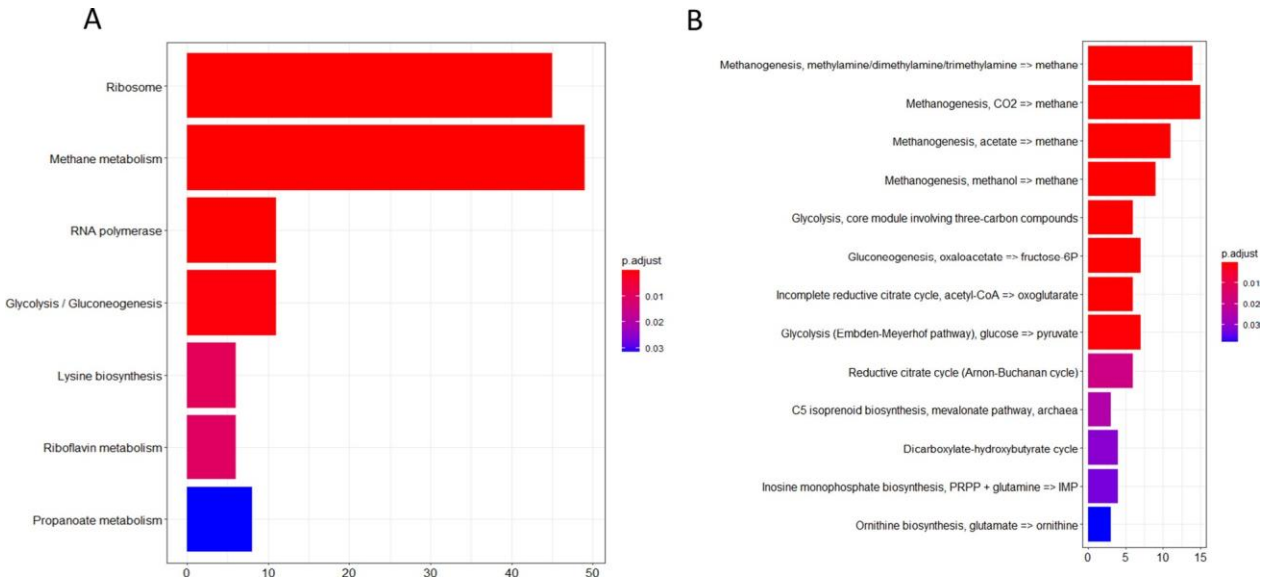


Figure 11.) Results of KEGG Pathway enrichment analysis (**A.**), and KEGG Module (**B.**). The pathways, which were significantly different between N2-MTR and H2-MTR samples are presented. X-axis indicates the number of KEGG IDs found as significantly different in the given pathway (listed along the Y axis). P-adjust stands for corrected p-values

5.2.6. Changes in the expression levels in methanogenesis genes

The enrichment analysis revealed that CH_4 metabolism was affected the most, hence I examined the changes in the expression levels of the MAG's genes involved in the various methanogenesis related metabolic pathways and modules in detail to gain a deeper insight

to the molecular mechanism behind. An overall of 154 genes of the 7 Archaea MAGs from the *Methane metabolism* (map00680) pathway were down-regulated (*log2FC* lower than -2), and 171 that were up-regulated (*log2FC* higher than 2), out of them 90 and 72 were significantly differentially expressed based on the *p*-value threshold of 0.05. Figure 12 shows the *log2FC* of the genes of the KEGG map00680 pathway in each MAG that harbored more than five of these genes. The two MAGs identified as belonging in the genus *Methanobacterium* (bin_35 and bin_51) and *Methanoculleus* (bin_6, bin_60 and bin_66) showed a very similar response (Figure 12.), many of their *map00680* genes were expressed at *log2FC* higher than 2, i.e., four-times higher expression. Two additional *Methanoculleus* MAGs (bin_60 and bin_66), a low and a medium quality MAG (according to CheckM estimation), were identified but not presented in Figure 12. This implies that several *Methanoculleus* strains actively participate in the BBU reaction.

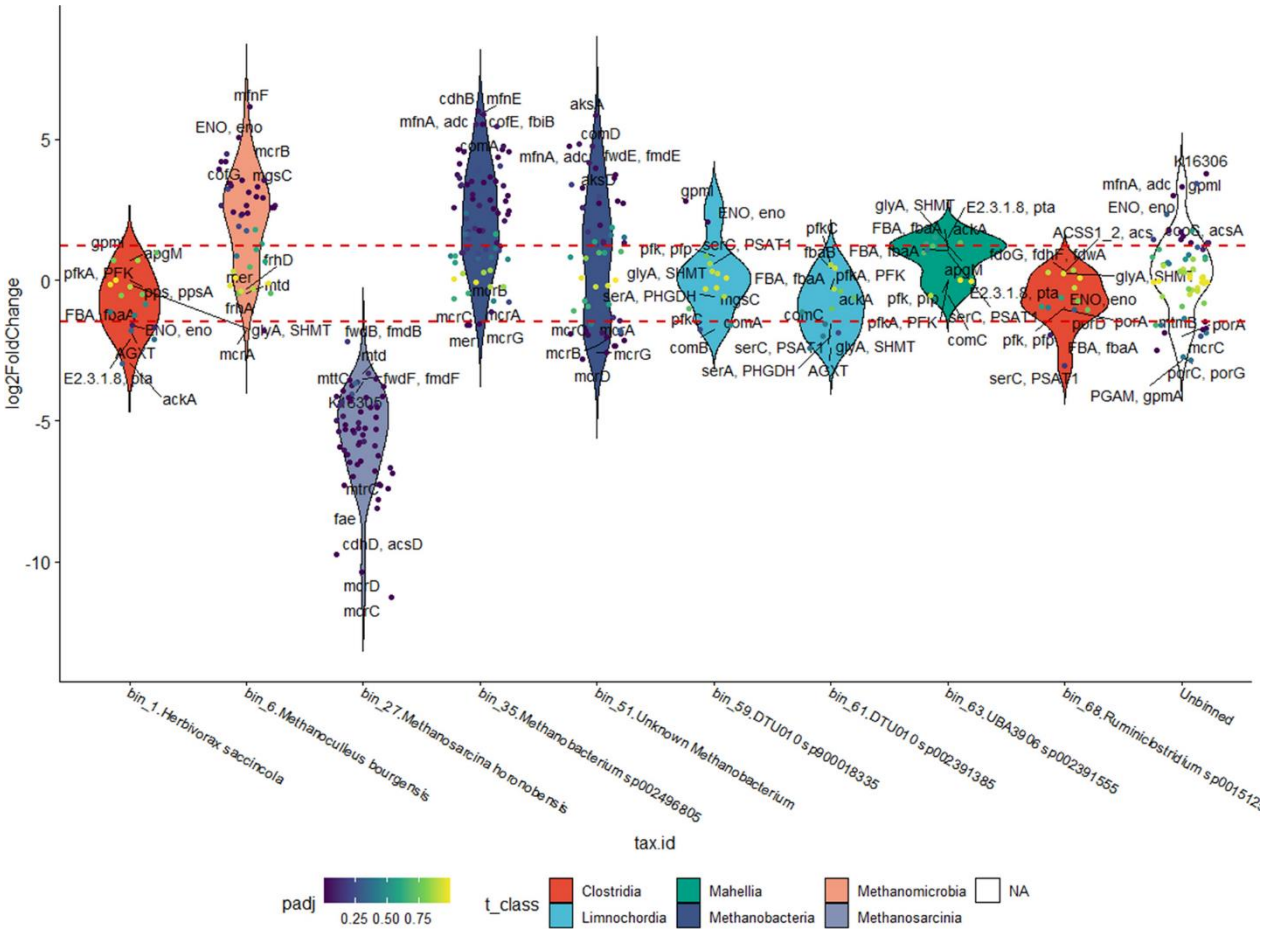


Figure 12.) Violin plot of genes (dots) involved in the methanogenesis KEGG pathway (map00680) in each MAG (arranged on the X-axis) and the unbinned gene collection. Only MAGs, which contain at least 5 methanogenesis genes are plotted. Filling colors indicate taxonomy at Class level. Each dot represents a KEGG orthologue (KO) in the respective MAG. Colors of the dots indicate the p-value of the *log2FC* difference between N2-MTR and H2-MTR samples. Horizontal dashed red lines mark the *log2FC* thresholds for significantly different KOs (respective $p\text{-}value < 0.05$).

The expression level of numerous genes increased shortly after H_2 injection in the hydrogenotrophic strains, which indicated that several metabolic pathways responded to the increased H_2 concentrations. The *log2FC* values of the genes ENO (phosphopyruvate hydratase, EC 4.2.1.11), COF (7,8-didemethyl-8-hydroxy5-deazariboflavin synthase, EC 4.3.1.32), and COM (sulfopyruvate decarboxylase, EC 4.1.1.79) were the largest in *M. bourgensis*: 5.06, 4.2 and 1.25, respectively. The ENO enzyme takes part in the biosynthesis of Coenzyme B, which is an essential molecule in the final step of the methanogenesis. The COF enzymes are responsible for the synthesis of the other important coenzyme, Coenzyme F420. The COM enzymes catalyze the 3-sulfopyruvate to 2-sulfoacetaldehyde reaction, which is an intermediate step in the synthesis of the third important coenzyme, Coenzyme-M (Grochowski and White, 2010). These results clearly suggested that the cells quickly started to increase the synthesis of all coenzymes, which were involved in methanogenesis to support the quick conversion of H_2 and CO_2 to CH_4.

The hydrogenotrophic methanogenic cells activated many key enzymes in the methanogenesis pathway to consume more effectively the H_2 from the environment. For example, in the two MAGs belonging *Methanobacterium* strains, the expression level of the KEGG orthologues associated with the enzymes MFN/ADC (tyrosine decarboxylase / aspartate 1-decarboxylase), FMD/FWD (formylmethanofuran dehydrogenase), AKS (methanogen homocitrate synthase) and COM increased significantly. These enzymes play an important role in the hydrogenotrophic methanogenesis pathway. The MFN and ADC enzymes are normally involved in methanofuran biosynthesis, while the FMD/FWD redox enzyme complex contains a molybdopterin cofactor and numerous [4Fe-4S] clusters to catalyze the reversible methanofuran to formyl-methanofuran reaction (which is an

important step in CO_2 conversion and in the oxidation of coenzyme-M to CO_2). The AKS enzyme takes part in the synthesis of Coenzyme-B. It is noteworthy however, that MCR enzymes (methyl-coenzyme-M reductase), which catalyze the final step of the methanogenesis showed lower expression in all hydrogenotrophic MAGs (Supplementary Figure S3). A likely explanation for this observation is that two hours was not enough for redirecting this section of the methanogenesis pathways. If the local substrate availability did not increase significantly, the cells did not need to increase the transcriptional activity of the MCR enzymes.

On the other hand, almost all genes in *Methanosarcinia honorobensis* showed a significantly decreased expression in the presence of H_2. This strain has been described as acetotrophic, but showed to also grow on methanol, acetate and various substrates but not on H_2/CO_2, formate and others (Shimizu et al., 2011). Its expression levels of MCR, ACS (acetyl-CoA decarbonylase/synthase) and FAE (5,6,7,8-tetrahydromethanopterin hydro-lyase) were significantly decreased. The ACS enzyme is responsible for the conversion of acetate to acetyl-CoA, which is a typical step in the acetotrophic methanogenesis pathway. The subsequent enzyme, FAE generates 5,10-methylene tetrahydromethanopterin (5,10-Methylene-THMPT) from formaldehyde, an important intermediate of methanogenesis. The substantial decrease in the transcriptional response of *M. honorobensis* to H_2 injection corroborated that this strain is unable to utilize H_2 and signaled an active inhibitory role of H_2 on acetotrophic methanogenesis. This implicates a hitherto unrecognized tight regulatory role of H_2 on diverse pathways coupled to methanogenesis.

5.2.7. Interactions between methanogenesis and other metabolic processes in Archaea

In addition to the methanogenesis pathways in the Archaeal MAGs, I identified nine pathways that were expressed differently as the early response of the microbiota to H_2 injection. Figure 13. presents the Archaea and Bacteria MAGs that indicate substantial up- or down regulation of several KEGG pathways. H_2 addition rapidly caused gene expression

changes in the Archaea, i.e., bin_6, bin_27, bin_35 and bin_51, since the *Ribosome, RNA polymerase* and *Methanogenesis* pathways were altered mainly in these MAGs.

In the case of Archaea, one *Methanoculleus* MAG (bin_6) and the two *Methanobacteria* MAGs (bin_35 and bin_51) responded with elevated gene expression in all pathways, while the *Methanosarcinia* (bin_27) and *Iainarchaeia* (bin_18) responded with a substantial and general loss of transcripts, i.e., biological activity, in them. Interestingly, the three *Methanoculleus* MAGs responded differently to the H_2 injection. Apparently, the entire metabolic activity, including all KEGG orthologs, were tuned up in bin_6 (classified as *M. bourgensis*), whereas only *Ribosomal activity, RNA transport* and *Lysine biosynthesis* was strongly upregulated in bin_60 and hardly any change in metabolic activity took place in bin_66 representing presumably a separate strain of *M. bourgensis*. Their overall gene expression did increase (*log2FC* of 2.19 and 2.37, respectively), thus the observed differences by bin_60 and bin_66 might as well indicate a slower response or perhaps further H_2 addition would have triggered a response more similar to that of the abundant *M. bourgensis* (bin_6). This observation may indicate the time resolution limit of H_2 triggered transcription and metabolic changes. Nonetheless it seems that the whole RNA machinery must be altered for responding to a significant change in the environment. Indeed, almost all genes (including the subunits of RNA polymerase for instance) from these pathways were highly expressed in the *Methanomicrobia* and *Methanobacteria* MAGs, and 64% of them with a *log2FC* of 2 or higher (and with a *p*-value of 0.05 or lower). The response to H_2 injection was quite the opposite in the case of *Methanosarcina horonobensis* (bin_27), as the expression of all investigated KEGG orthologs and metabolic pathways were hindered significantly, i.e., up to 33% (Figure 13.).

Other carbon metabolism-related pathways that showed an overall significant difference in the PWE analysis were *Carbon fixation* and *Glycolysis/gluconeogenesis* – these showed a similar pattern. For example the *folD* gene of the reductive acetyl-CoA pathway (Wood-Ljungdahl pathway) was transcribed vigorously in bin_6 (*M. bourgensis*) (*log2FC* = 3.7). The relative enrichment of "*Methanogenesis, acetate to methane*" was overall the highest in this MAG (mean *log2FC* = 3.55), this can be linked to the elevated acetotrophic

methanogenesis, as there was no other major difference between the expression change in these pathways. Interestingly though, the *Methanogenesis, CO_2 to methane* module did not increase drastically (nor did the methylotrophic module), with the exception of a handful of genes showing *log2FC* higher than 2, including *methenyl-tetrahydromethanopterin cyclohydrolase* gene in bin_6 and bin_35 (*log2FC* = 2.56 and 3.49, respectively), and some others with smaller but still significant differences, including the *F420-non-reducing hydrogenase iron-sulfur subunit* gene of bin_6 (*log2FC* = 1.32, p-value = 0.04).

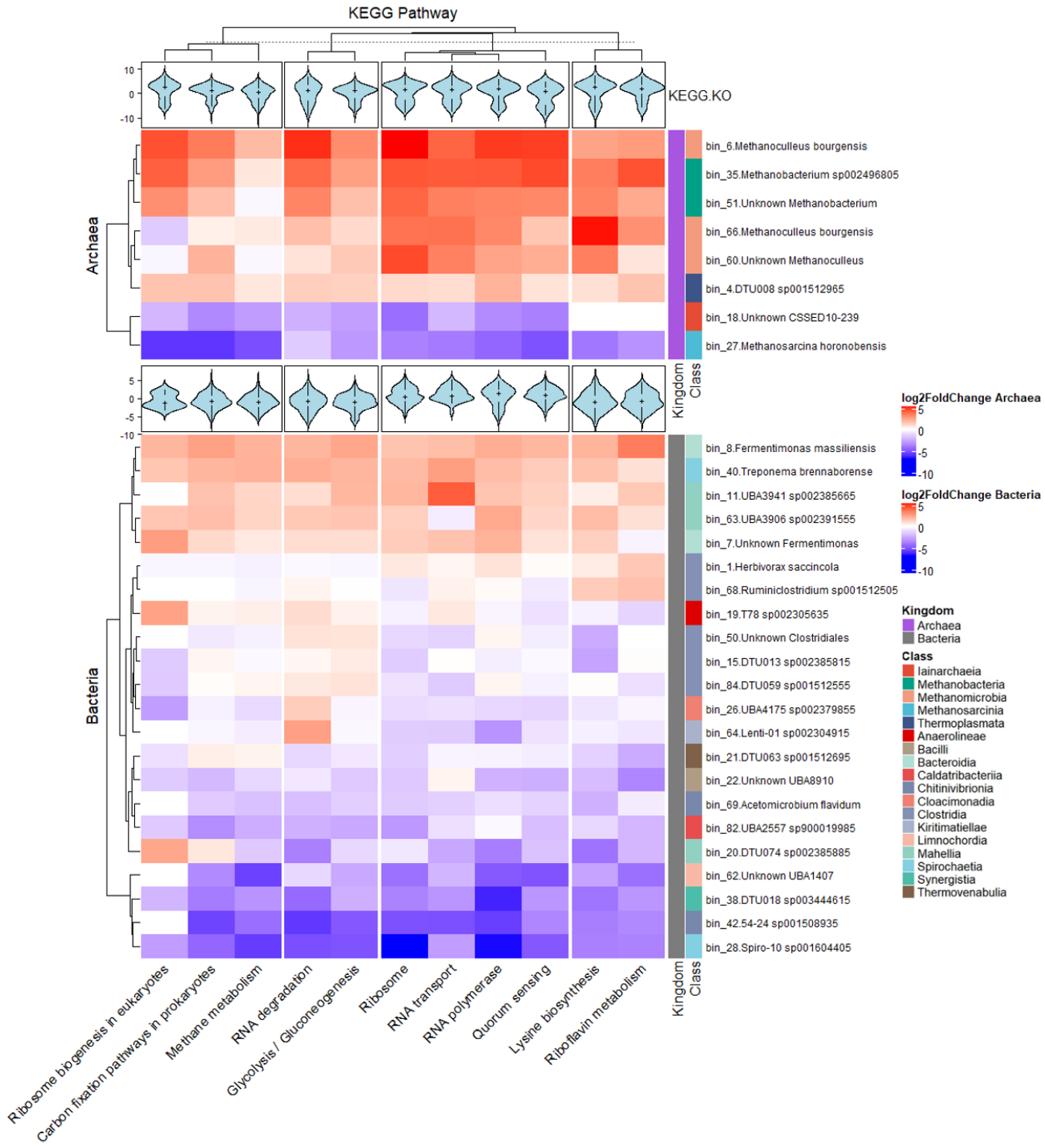

KEGG Pathway
KEGG.KO
Archaea
Bacteria
bin_6.Methanoculleus bourgensis
bin_35.Methanobacterium sp002496805
bin_51.Unknown Methanobacterium
bin_66.Methanoculleus bourgensis
bin_60.Unknown Methanoculleus
bin_4.DTU008 sp001512965
bin_18.Unknown CSSED10-239
bin_27.Methanosarcina horonobensis
bin_8.Fermentimonas massiliensis
bin_40.Treponema brennaborense
bin_11.UBA3941 sp002385665
bin_63.UBA3906 sp002391555
bin_7.Unknown Fermentimonas
bin_1.Herbivorax saccincola
bin_68.Ruminiclostridium sp001512505
bin_19.T78 sp002305635
bin_50.Unknown Clostridiales
bin_15.DTU013 sp002385815
bin_84.DTU059 sp001512555
bin_26.UBA4175 sp002379855
bin_64.Lenti-01 sp002304915
bin_21.DTU063 sp001512695
bin_22.Unknown UBA8910
bin_69.Acetomicrobium flavidum
bin_82.UBA2557 sp900019985
bin_20.DTU074 sp002385885
bin_62.Unknown UBA1407
bin_38.DTU018 sp003444615
bin_42.54-24 sp001508935
bin_28.Spiro-10 sp001604405
Kingdom
Class
Ribosome biogenesis in eukaryotes
Carbon fixation pathways in prokaryotes
Methane metabolism
RNA degradation
Glycolysis / Gluconeogenesis
Ribosome
RNA transport
RNA polymerase
Quorum sensing
Lysine biosynthesis
Riboflavin metabolism
log2FoldChange Archaea
log2FoldChange Bacteria
Kingdom
Archaea
Bacteria
Class
Iainarchaeia
Methanobacteria
Methanomicrobia
Methanosarcinia
Thermoplasmata
Anaerolineae
Bacilli
Bacteroidia
Caldatribacteriia
Chitinivibrionia
Cloacimonadia
Clostridia
Kiritimatiellae
Limnochordia
Mahellia
Spirochaetia
Synergistia
Thermovenabulia

Figure 13.) Heatmap of significantly various *KEGG Pathways* in MAGs that harbor at least 10 genes in the respective pathway. Top panel shows Archaea, while the bottom panel shows Bacteria MAGs. Filling colors are scaled to the *log2FC* of all the genes in the respective pathway of the MAG. Violin plots at the top show the *log2FC* values of every gene participating in the respective pathway.

5.2.8. Changes in gene expression levels in bacterial MAGs

Some genes involved in, or related to elements of the methanogenesis pathway could be found in bacterial MAGs as well, e.g. in *Herbivorax saccincola, Ruminiclostridium sp001512505*, in two unknown *Limnochorida* and in one *Mahellia* MAG. However, when inspecting the change of the methanogenesis-related KEGG orthologs in these MAGs, it became clear that these genes showed significant difference only in a few cases, i.e., their *log2FC* values were spread between the threshold lines that indicated significance. Consequently, they were involved in the overall methanogenesis, and related metabolic pathways (which are included in the KEGG map00680 pathway), but they did not respond to the H_2 stress. This was a substantially different behavior from the one that the Archaea MAGs showed, which clearly expressed their genes differently as a respond to H_2 injection.

In the case of Bacteria, the RNA-machinery pathways (ko03010) showed an overall decrease in gene expression, except for bin_40 (*Treponema brennaborense*), bin_8 (*Fermentimonas massiliensis*), bin_11 (*UBA3941_ sp002385665*) and bin_7 (*Unknown Fermentimonas*). These MAGs had low abundance, though they showed an increase in the MTR samples. The related pathways seem to be up-regulated in bin_40 and in bin_11 (mapped in class *Mahellia*, order *Caldicoprobacterales*). Most of the small and large ribosomal subunits showed *log2FC* of 2 or higher. Another member of the family *Treponemataceae* (*bin_28 Spiro-10 sp001604405*) showed a clear downregulation in all discussed pathways.

In the AD process, *Treponema* behave like homoacetogenes, they consume H_2 and CO_2 to produce acetate, hence they may compete with hydrogenotrophic methanogens

(Kotsyurbenko et al., 2001), although not very efficiently (Wahid et al., 2019). I identified only two methanogenesis related genes in bin_28 and bin_40 (formate-tetrahydrofolate ligase and methylenetetrahydrofolate reductase NADPH), bin_40 showed an overall activity increase (*log2FC* = 2.216), indicating either that this pathway would become more active at a later timepoint, or these bacteria utilize alternative catabolic activities. In a relevant observation *Treponema* abundance increased in digesters spiked with H_2 (L. Li et al., 2016) although after 90 h the signs of H_2 stress were noted in the digester. Essential genes of the Wood-Ljungdahl (WL) pathway were apparently not expressed in bacterial MAGs in a recent study (Zhu et al., 2020). In contrast, in the present work we identified several MAGs harboring these genes, including bin_7 (Unknown *Fermentimonas*), bin_8 (*Fermentimonas massiliensis*) and bin_20 (DTU074 sp002385885, although all of them showed low abundance (~ 0.3–1%). Interestingly, bin_20 exhibited an overall decrease, but the expression of its WL pathway genes increased. This can be attributed to the elevation of the transcriptional activity of only two genes, *fhs* (formate-tetrahydrofolate ligase) and *folD* (methylene-tetrahydrofolate oxidase), which are important in WL pathway (*log2FC* = 6.31 and 3.14, respectively). This response to H_2 is thus the opposite to that of bin_40, suggesting that as acetogenic methanogenesis increased, this bacterium might have tried to compete with the Archaea for acetate. The other two homoacetogens, which increased their transcriptional activity (*log2FC* = 1.40 and 2.56, respectively), apparently included the *fhs* and *folD* genes as well. It was also demonstrated earlier that homoacetogenic microbes tended to increase their activity in a H_2-fed systems (L. Treu et al., 2018).

6. Discussion

One of the major aims during the course of my doctoral work, was to investigate the possibility of utilizing GWB as a feedstock for biogas digesters. Using GWB would have numerous advantages, including its low costs and the option to grow in marginal lands where it would not come into contradiction with crops grown for food/feed purposes (Valentine et al., 2012). To this end, I harvested plants in 3 occasions during the growth season (June, August, October) from a small research plantation, consisting of 100 plants from 2 different genotypes, EN and PP (Dudits et al., 2016). I compared both the leaves and the stems and the WWB's chemical (fiber) properties, fermentation parameters and microbial compositions in batch anaerobic reactors. I used the Gompertz model to estimate fermentation rates and final yields (μmax and K values, respectively). I analyzed the leaves and stems separately (Table 1., Figure 1.) but also in a mixture, wherein their ratio was equal to what the plants themselves showed in each harvest (L/S ratio), as the biomass from the harvesting of these plantations on an industrial scale would be more similar to this latter arrangement (Table 2., Figure 2.).

My results supported that in both GWB and WWB the lignin concentration negatively correlated with fermentation rates and final yields. This relationship had been already established for several other substrates (Herrmann et al., 2016). The solubles part, however, correlated positively with these parameters. When comparing the three harvests, I found that there was a substantial difference in the plant's μmax values and to a lesser extent in their K values as well. The June harvest showed the highest values: it was significantly higher than the other two in both the separate and in the mixed L/S fermentations. This was probably the consequence of the higher leaf biomass (higher L/S ratios) in these harvests. The results of the separated and mixed L/S fermentations showed minor differences. More interestingly, PP plants showed better K and μmax values than their parent strain, EN, in almost every comparison.

Lignin is resilient against anaerobic biodegradation and thus is the main target of pretreatments utilized to improve the efficiency of the AD of lignocellulosic substrates

(Isroi et al., 2011; López et al., 2013; Schroyen et al., 2015). Thus, pretreating this lignocellulosic substrate may yield superior fermentation performances. A combined fungal and alkaline pretreatment was used to improve the anaerobic digestion of woody willow substrates before (Alexandropoulou et al., 2017) (there termed willow sawdust, WSD), this approach may yield decisive results in the case of GWB as well.

Another possibility is to breed energy crops with lowered lignin content. These efforts are currently underway (Allwright and Taylor, 2016; Cho et al., 2021; Dudits et al., 2022). Lignin confers rigidity to cell walls, it is associated with the mechanical stability of plants and the tolerance against abiotic and biotic stresses (Frei, 2013), thus variants with too low lignin content may show lower plant fitness and biomass yields (Bonawitz and Chapple, 2013; Littlewood et al., 2014) and lead to overall lower bioenergy potential. For example, in a field trial with genetically modified poplars the down-regulation of cinnamoyl-CoA reductase (CCR), an enzyme in the lignin biosynthetic pathway, resulted in yield penalty in many of the transgenic lines, which if taken into consideration, caused a lower final ethanol yield.

In an attempt to characterize the effect of cell wall composition and accessibility of glucan of several SRC willow genotypes to enzymatic saccharification via fungal cellulases after a dilute acid pretreatment (Ray et al., 2012) found that higher lignin content of the biomass was not necessarily detrimental to final glucose yields and projected ethanol yields. Albeit aerobic cellulose degradation by fungal cellulases is different from anaerobic cellulose degradation by the complex microbial consortium in AD. Moreover, as there are conflicting reports on both biomass yield and downstream processing efficiency resulting from the engineering of lignin pathway genes, most likely these parameters are highly influenced by environmental factors as well (Chanoca et al., 2019).

Breeding of plants with altered lignin composition seem to be a possible solution to this dilemma. However, in a recent metastudy wherein 21 lignin biosynthesis genes were perturbed in *P. trichocarpa*, and the results of transcriptomic, proteomic, fluxomic, and phenomic data of 221 lines were comprehensively integrated, the authors concluded that

tree growth is not associated with lignin amount, subunit composition or specific linkages, but rather correlated with the presence of collapsed xylem vessels, the activation of a cell wall integrity pathway and/or the accumulation of chemical inhibitors (Wang et al., 2018).

A promising solution is the activation of the heterologous expression of microbial lignin-degrading enzymes *in planta* at late stages of plant growth. A bacterial lignin degrading enzyme was accumulated in its active form, activated *in planta*, and used to improve biomass saccharification efficiency of *N. benthamiana* (Ligaba-Osena et al., 2017). Lignin-degrading enzymes of *Basidiomycete* origin were also expressed via an inducible strategy (Khlystov et al., 2021). The authors concluded that this strategy was advantageous as plant fitness effects would be mitigated relative to existing strategies (Chanoca et al., 2019), which aimed at reducing and/or modifying lignin biosynthesis. If this approach could be adapted to poplar and willow, it could result in a tremendous amount of available lignocellulosic biomass ready for harnessing.

I also carried out a 16S rRNA gene-based metataxonomic analyses and the results showed a dominance of class *Clostridiales*. Generally, the high proportion of *Clostridiales* in AD is not unusual, for example in a thermal hydrolysis process (THP) pre-treated of wastewater sludge and household waste anaerobic co-digestion experiment up to 68% abundance of *Firmicutes* (composed of mainly *Clostridiales*) and a *Firmicutes:Bacteroides* ratio of 3-5 was observed (Westerholm et al., 2019). Here, the *Firmicutes*/*Bacteriodes* ratio positively correlated with the CH_4 yield. An even higher ratio (5–6 was observed by Güllert et al. (Güllert et al., 2016), although the authors concluded that the low abundance of *Bacteroides* in the AD sludge was a sign of a less effective plant material decomposition compared to herbivore guts. The central role of *Clostridiales* was noted several times before, they are essential for the proper functioning of biogas digesters (Kougias et al., 2018; Wirth et al., 2012). The relative abundance of these microbes is normally within the 40–60% range (Wirth et al., 2012), but this can undoubtedly be influenced by the used substrate and fermentation conditions. Several species of *Clostridiales* are known for their cellulolytic properties, due to the presence of genes encoding the multi-enzyme complex cellulosome (Koeck et al., 2016) and other hydrolytic enzymes. Nevertheless, it is unusual

that the phylum *Bacteroides* is essentially missing from the AD microbial community (only 3 ASVs from this taxon were found and their abundance was below 0.1%).

Industrial biogas plants require large amounts of substrates daily, thus when considering the feasibility of novel substrates, the biomass yields per area should be considered (Kakuk et al., 2021a). The Energo GWB yields were 22.8 and 31.5 t VS/ha in the August harvest and in the September estimation, respectively. These are significantly higher than that of described in a 10-year-long study conducted in Poland — these ranged from 2.2 to 22.5, with an average of 13.1 t/ha (Stolarski et al., 2017). This is probably because of the leaf biomass, which in my results contributed to 1.22 times the mass of the stems (Table 2: June harvest, PP). Moreover, if we consider energy potentials (calculated from the energy content of methane yield and from the net calorific values when combusting the biomass), the biomass utilization methods can be compared. Considering this, I carried out an estimation of per hectare energy yields from the three GWB harvest data and compared it to WWB and maize silage. This showed that GWB can be a promising substrate for biogas generation. The Ep/ha values from the August harvest and September estimation were 176.27 and 227.65 GJ/ha/year, respectively, which are in the range of maize silage and are higher than that of WWB when it is used in anaerobic digesters, and similar to WWB when it is used for burning. It should be emphasized though that compared to WWB, GWB performs better from an industrial-scale point of view partly because of its higher Ep, and partly due to its higher μmax values. Faster CH_4 production is more advantageous in industrial CSTR reactors.

Regarding optimal harvesting time, I found that the highest biomass yields should be expected in August-September. In a study conducted on *Miscanthus* harvesting time had a large impact on the performance of AD unlike on biomass combustion. The average AD energy potential differences over all harvest points was 35% less than for combustion. However, if optimal harvesting times are taken into account, which is autumn green harvest for AD and a delayed harvest for combustion, the related energy loss can be decreased to an average of 18% (Kiesel et al., 2017). Overall, the authors concluded that anaerobic digestion is a promising utilization pathway for miscanthus biomass. Moreover, biogas is a

higher-value energy carrier than heat, as it can be upgraded to biomethane, introduced into the existing natural gas infrastructure and further utilized as a transport fuel. AD technology also offers the additional benefit of easy nutrient recycling via digestate relative to ash from combustion.

The possible attenuation of annual cutbacks on the plants should be taken into account when calculating economic feasibility. The authors of a review of willow production in Poland (Stolarski et al., 2019) compared different harvesting frequencies, among others, and found that the annual harvesting of the woody biomass was comparable to the other scenarios (biannual, triannual and quadrennial harvests) in biomass yields. It has been noted that the plantations showed very high variances caused by numerous factors, such as soil type, cultivar and agro-technologies used (planting density, fertilization, etc.). Thus, it is unlikely that the annual cutbacks would have a significant effect on the biomass yields in GWB.

Another important consideration is to develop a scheme for the storage of GWB, to supply biogas plants with fresh feedstock all-year-around. In case of dry WWB, organic matter losses during the storage can reach 20% (Whittaker et al., 2018), while a good management can lead to as small as 1% dry material loss upon ensilaging (Borreani et al., 2018). Considering this and given the obvious resemblance of GWB to other plant materials that are frequently ensiled, a plausible approach for the preservation of GWB can be ensiling. Indeed, preliminary results on ensiling GWB supported this approach (Nyári et al., 2021).

The second major aim during the course of my doctoral work was to investigate the behavior of the AD microbial community and in particular the hydrogenotrophic methanogens in a BBU system via genome-centric metatranscriptomics. Currently the overwhelming majority of renewable energy production approaches are photovoltaics and wind power. While readily available and technologically mature, both generate electricity in an intermittent fashion. As the power distribution grids are designed to balance electricity production and consumption continuously, they can operate in a fluctuating

mode only with a substantial energy loss. Hence, technologies to balance the fluctuations are urgently needed. A very promising solution to this problem is offered by the flexible anaerobic digestion technology (Liebetrau et al., 2020), as biogas plants have controllable energy output to buffer the fluctuations in renewable electricity production. Moreover, with a coupled technology called Power-to-Biomethane (P2bioCH_4), biogas reactors can efficiently convert the temporarily surplus renewable electricity to biomethane (bioCH_4) in a carbon neutral or negative carbon footprint process way. Clean bioCH_4 is chemically indistinguishable from the fossil natural gas, therefore it can be stored and transported efficiently and inexpensively in the natural gas grids. The biotechnological route to P2bioCH_4 is carried out via biological biogas upgrading involving methanogenic Archaea, crucial members of the AD microflora, a group of rare and obscure obligate anaerobic microbes. Among them, the hydrogenotrophic methanogens convert H_2/CO_2 into CH_4, but our knowledge regarding the enzymatic processes leading to and regulating CH_4 formation are far from complete. Understanding of the molecular regulation and control of the highly complex cell factory pathways of microbial communities carrying out AD, is a challenge for both basic and applied research. The advantages offered by P2bioCH_4 technology is to convert bursts of H_2 resulting from intermittent water electrolysis using excess electricity that is generated by photovoltaics or wind energies (Kakuk et al., 2021b; Szuhaj et al., 2016).

Metagenomic (MG) and metatranscriptomic (MTR) methods help illuminating the workings of biogas producing communities (Jünemann et al., 2017). Campanaro et al. used the metagenomes from 134 different published sources to assemble to the *biogas microbiome* (Campanaro et al., 2020), and estimated the abundance of each MAG in those samples. This revealed a very high plasticity of the AD microbiome and the presence of multiple different microbial communities that have little to no overlap among them. The occurrences of these MAGs in the samples and number of MAGs in each of the metagenomes are shown in Supplementary Figure S3. On average 757 MAGs were found (with higher than 0.00001% abundance) in the metagenomes. This showed that several MAGs are unique to different fermentors/fermentation parameters, and also that some

members of the microbiome in those metagenomes, where only a low number of MAGs were found, could not be assembled and thus analyzed. This urges to further extend this procedure to other fermentation samples and conditions.

In a recent study, thermophilic biogas reactors were fed with H_2, and after 18 h and 36 days MTR analyses were carried out to unveil the involvement of the individual MAGs in the global microbiome functions (Zhu et al., 2020). Another study, which was the first that investigated biogas reactor metatranscriptome dynamics following hydrogen injection used a one week-difference (Fontana et al., 2018). Thus, the effects of H_2 addition on the metatranscriptome of a thermophilic AD in mid- and long-term were studied before. However, to exploit the potential of the P2bioCH_4 technology to the most, the short-term response of the microbial community, and in particular the hydrogenotrophic methanogens to these bursts of H_2 must be studied in detail as well. While the scientific literature mainly focused on thermophilic reactor configurations, many industrial biogas reactors operate at mesophilic temperatures. Combining genome-resolved metagenomics with metatranscriptomics, the analysis of the total mRNA content of an environmental sample, can yield even deeper understandings of the participants of the biogas producing consortium and their genetic activity. This shows the importance of using genome-resolved metagenomic analyses in AD microbiome research.

For these and the reasons explained in the *Introduction section*, I aimed at identifying the early response of a mesophilic microbial community to this environmental change, with a particular attention to methanogens, a scenario frequently envisaged and expected in the P2bioCH_4 industry (Kakuk et al., 2021b). The investigation of H_2 stress effect on a complex anaerobic metagenome in such a high time-resolution via metatranscriptomics has not been carried out yet. Therefore, the results presented here may contribute to a general knowledge in this regard as well.

According to my experimental setup, an alpha-cellulose-adapted AD community was saturated with H_2 gas for 10 minutes, and samples taken before, and immediately after the

treatment were used for community DNA and RNA sequencing. In order to gain higher statistical confidence in the results, this process was repeated after 2 months. During this experiment firstly I established that the composition of the microbial community at DNA-level did not change significantly, as the different reproduction rates of the various microbes would have disturbed the picture of the early functional response. The practically identical results obtained (Figure 2.) from the four metagenome (MG) sequencing datasets confirmed that this indeed was the case, and furthermore supported the reproducibility of the system. Thus, the combination of the four MG sequencing libraries were used to assemble a fairly large number of MAGs: 7 Archaea, 61 Bacteria and 16 unclassified population genomes. My analysis showed that from these, 14 had no corresponding genomes in the BGMB database. I concluded that although the microbial communities in our reactors showed similarities to other biogas samples, certain microbial taxa were characteristic to this system only. The high number of unique MAGs in each anaerobic digester sample was suggested in the related comprehensive analysis of microbial genomes (Campanaro et al., 2020). My results here support the importance of using genome-resolved metagenomics to analyze biogas producing microflorae, and also complex microbial communities in general.

The community switched to H_2 consumption and biomethane production almost immediately following H_2 injection, although feeding of the entire community with alpha-cellulose substrate continued as before. Our research group previously demonstrated that the microbial community composition rearranges upon long-term exposure to H_2 (and CO_2), particularly when no other organic substrate is available for the community (Ács et al., 2019). The results of the current work indicated the presence of sufficient hydrogenotrophic methanogenesis activity to perform the P2bioCH_4 conversion rapidly in the industrial biogas plant effluent. In other words, the diverse, “raw” anaerobic communities can be used in switching on P2bioCH_4 without a lengthy adaptation and enrichment period. This allows a quick and efficient turn-on and turn-off response by the mixed methanogenic community. The vigorous P2bioCH_4 activity returned to normal

biogas production as soon as the dissolved H_2 diminished, but the community was ready to adjust its biochemistry to instant H_2 conversion and P2bioCH4 repeatedly.

I examined the initial response of the complex AD microbial community with a custom metatranscriptomic bioinformatics workflow to assess the first up- or down-regulated genes caused by the H_2 injection. I found that the methane metabolism in the *Methanosarcina* MAG (bin_27) practically stopped (average *log2FC*= -5.23); also in an unknown *Iainarchaeota* (bin_18) it decreased substantially (-2.85). Out of the two *Methanoculleus bourgensis* MAGs, one showed a significant methanogenesis pathway enrichment (bin_6), while the other indicated only a slight enrichment (bin_66), the average *log2FC* was 1.94 and 0.65, respectively. Another unknown *Methanoculleus* MAG (bin_60) was on the contrary somewhat decreased (-0.17). This supports that phylogenetically closely related microbial strains can behave differently in the same microbial environment, which also underlines the importance of using genome-resolved metagenomics and metatranscriptomics. In a recent study, where the effects of H_2-addition in thermophilic biogas digesters were examined (Zhu et al., 2020), the results revealed a multi-trophic role to *Methanosarcina thermophila*. Although the hydrogenotrophic *Methanoculleus thermophilus* prevailed as the dominant Archaea species in terms of relative gene expressions, at the expense of *M. thermophila*. Some community members, which emerged during the later stages of H_2-driven methanogenesis, were below the detection limit in the starting sample, i.e., *Methanobacteriaceae spp.* In my work a similar pattern started to emerge but in a mesophilic context.

In a brief review to summarize *in situ* biogas upgrading Zhang et al. pointed out the predominant roles of the genus *Methanoculleus* under mesophilic conditions and the thermophilic genus *Methanothermobacter* at elevated temperatures (Zhang et al., 2020). The species *Methanoculleus bourgensis* (bin_6 and bin_66 in my experiments) was identified to play an important role in various biogas reactor systems. *Methanoculleus* species grow on CO_2 and H_2 and hence perform the hydrogenotrophic pathway for CH_4 synthesis (Maus et al., 2015). The importance of this species was shown in my experiments as well.

H_2 (and dissolved CO_2) is readily converted to CH_4 by both direct (hydrogenotrophic) and indirect (homoacetogenesis and subsequent acetotrophic) methanogenesis. My findings suggest that the second route is unlikely the predominant one in the early response of the microbial community to H_2 addition at least under mesophilic conditions, since the acetotrophic pathways have not reacted significantly within the two hours after sampling, while the gene transcription of the hydrogenotrophic route increased dramatically after a very short period of extensive H_2 feeding (Figures 13. and 14. and Supplementary Figure S2). This predicts that under P2bioCH_4 operation conditions the physiological readiness of the hydrogenotrophic methanogens will determine the reactor response rate upon switch-on of the H_2 addition.

Interestingly, my experiments revealed an extensive response to the transient H_2 stress within the Bacteria community as well, although Bacteria cannot directly generate CH_4 from H_2 as many Archaea can. Some of these Bacteria possess the complete or partial enzyme sets for the Wood-Ljungdahl pathway. These and the homoacetogenes are probably the best candidates for syntrophic community interactions between members of the distinct kingdoms of Archaea and Bacteria. The details of these interactions in the complex anaerobic environment and consequences to stabilize robust and vigorous AD microbial communities during long term P2bioCH4 operation should be the subjects of future studies. Nevertheless, the transcriptional activity of the primary potential syntrophic bacterial partners (bin_1 (*Herbivorax saccincola*), bin_68 (*Ruminococcus sp.*), and unidentified bins_59, _61, _63, see Figure 14.) did not change substantially upon H_2 exposure. This may mean that either there is enough syntrophic capacity already in the non-adapted, “raw” community to support increased hydrogenotrophic methanogen activity or the syntrophic partners respond slowly to the sudden H_2 burst appearing in the microbial environment.

These results altogether emphasize the importance of carrying out the metagenomic binning procedures as opposed to be relied solely on databases when investigating complex environmental microbiota such as the AD microbiome. While the BGMB collection is a very useful database and necessary step in our more complete understanding of the AD

microbiome, further research is necessary to reach this goal, as novel microbial strains may exist in essentially every biogas reactor. Perhaps this is even more important in metatranscriptomic studies, as the mapping of reads to the actual population that is living in the sample (the MAGs) is more precise than when they are mapped to a database, as it is possible that closely related and less precisely annotated strains have different ORFs in the actual genome than in the database. This was suggested by Treu et al (Treu et al., 2018), as their results showed that some closely related genomes from key taxa for the specific bioprocess (e.g., *Methanoculleus* spp.) co-exist in the same ecosystem even though they possess different metabolic features. Although there is a caveat in this method as well, which is that there will be reads that do not map to the ORFs carried by the MAGs, but only the 'unbinned' fraction, as was the case in 16.34-20.97% of the reads in this study. A possible solution to this would be to map only these reads to the database or to increase the coverage of the metagenomes via deep sequencing.

Taking into account the recent results and considerations, the development of a stable P2bioCH_4 mixed AD community depends on a number of important parameters, such as the origin of inoculum, H_2 supply and its fluctuation, composition of added growth supporting substrates, the dissolved CO_2/HCO_{3-} concentration, temperature and reactor configuration.

In the future the extension of these studies should be carried out, i.e., mapping the molecular events after longer exposure of the microbial cell factory and linking the metagenomic approach to more detailed transcriptomic and proteomic studies.

The interest in converting the energy market from a fossil fuel-based one into one that is based more on renewables is growing worldwide. This positive trend will help avoid threatening global climate change and associated environmental catastrophes.

7. Conclusions

In the course of my doctoral thesis, I aimed to contribute to two important aspects of the anaerobic digestion technology: 1.) utilizing second-generation lignocellulosic substrates as feedstocks and 2.) biological biogas upgrading of biogas with H_2 resulting from the coupled process called Power-to-Gas. To achieve these goals, in the first part of my research, I evaluated a novel substrate, the juvenile stems and leaves of short rotation coppice willow, which I termed Green Willow Biomass (GWB), for biogas production. In the following section I will summarize my main findings:

The analyses of fermentation efficiency, based on parameters as maximal CH_4 yields (K) and highest CH_4 production rates (μmax), showed that both the commercial willow genotype (EN) and the novel tetraploid variant (PP) I tested can be viable alternative AD feedstocks:

1. The highest values can be expected from mid-summer harvests (June in my experiments)
2. Low lignin and high "solubles" are key factors for efficient AD of GWB, similarly to other previously characterized substrates.
3. Leaf biomass is beneficial, as its values were higher than the stem's.
4. PP is slightly better than EN in this context
5. Higher Leaf-to-Stem (L/S) ratios were shown to be positively correlated with μmax values, suggesting that PP plants harvested in June are optimal substrates for AD.
6. Compared to WWB, GWB has better fermentation parameters, even more so when considering CSTR fermentations and industrial-scales operations
7. In further research, breeding can improve biomass characteristics.

In addition to the laboratory experiments, I estimated the biomass yields (t/ha) based on a small-scale field experiment to compare GWB with alternative feedstocks. This showed that:

8. EN plants harvested in August produced VS amounts that was comparable to maize silage.
9. Compared to WWB the yields were higher because the large amount of leaf biomass.
10. An extrapolation suggested that September harvest could result in even higher yields.
11. If we consider fermentation parameters and biomass yields, a late August-early September can be the choice for harvest. This is because the extractable energy potential (GJ/ha) in August – September harvests of GWB reaches and can even surpass the corresponding yields of either maize silage or combusted WWB.

Energy willow shrub plantations, besides other advantages, can be harvested annually thereby reducing biomass production costs considerably. The annual cutbacks likely won't affect substantially the yearly biomass production. Ensiling may help to store and/or improve the quality of the green biomass for AD. Further studies, spanning several years and hectares are required however to validate the possibility of this scenario, e.g. annually harvesting the GWB substrate, ensiling it and using it CSTR reactors.

These findings help optimizing the GWB-based technology by using different genotypes, harvesting time, different lignocellulose composition and anaerobic microbial community. Since cultivating willow biomass has several advantages both from an economical and from an ecological point of view, anaerobic digestion of GWB has the potential to produce biogas for a lower cost and with more benefits than many other feedstocks. Therefore, the presented data and calculations provide substantial support to promote and recommend green willow biomass (GWB) as valuable feedstock for biomethanation.

Within the course of the second part of my thesis, I examined the behavior of the AD microbial community in a suggested P2bioCH_4 process on the DNA and mRNA levels. More precisely, I examined the early response of the mixed biogas microbial community to the presence of saturating amount of H_2. I carried out metagenomic and metatranscriptomic analyses to determine the changes of the expression levels of the various genes of assembled MAGs, with a focus to those related to methanogenesis. Genome-centric metatranscriptomics could expand our current understanding about the early response of the microbial community and thus contribute to manage the turn-on and turn-off steps of the P2bioCH_4 process. These investigations enabled the distinction of the activity of each individual MAG and the identification of the key and most sensitive members of the community. My main findings were the following:

1. The microbial community responded practically instantaneously (2 hours after the induced H_2) to the changed environment.
2. Many MAGs that have no correspondent in the biogas microbiome database are not always just side-players, rather several of them are abundant, main-players of the community. Thus, it seems that related, but specific microbes proliferate in the samples, which is likely the case in ultimately any AD environment. This underlines the importance of using genome-centric metagenomics for the analyses of complex microbiomes.
3. The activity of pluripotential methanogens, mainly *Methanosarcina* MAGs, reduced significantly, while the hydrogenotrophic *Methanobacter* MAGs increased.
4. In addition, the metabolic activity of numerous bacterial strains changed substantially as a response to H_2.
5. Clearly, the excess H_2 does not only affect the methanogenesis pathways in Archaea, rather the microbial community respond with a multifarious gene expression profile change, which seems to be rather selective. This indicates a more global regulatory role of H_2 in the life of anaerobic communities than assumed earlier.
6. The syntrophic interactions contribute to the stability and metabolic activity of the hydrogenotrophic methanogens.

7. This, together with the non-sterile operation conditions and continuous supply of inexpensive catalyst, underlines the benefits of using mixed communities in the P2bioCH_4 process instead of pure hydrogenotrophic cultures.
8. The metatranscriptomic responses to the H_2 treatments separated 2 months apart were very similar to each other indicating not just the robustness of the employed methods, but also that the metabolic pathways could be flexibly restored after switching on and off the P2bioCH_4 operational mode.

This thorough analysis of the differences between the H_2-treated metatranscriptomes and corresponding controls identified the early events in the microbial communities brought about by the H_2 addition, contributing to our understanding of the BBU mechanism in a P2bioCH4 system.

8. Acknowledgements

Throughout my work at the Department of Biotechnology and during the writing of this dissertation, I have received a great deal of support and assistance.

First and foremost, I would like to express my gratitude to Professor Kornél L. Kovács, who was my main supervisor and mentor during my work conducted at the Department of Biotechnology, from when I was only an undergraduate worker to this day. His knowledge about and passion in biotechnology encouraged me to take biotechnology as a career path as well. He always supported me and my work, my ideas, and he let me follow my own ways even when they were not perfect. Without his guidance, I could not have become a researcher.

I also owe a big thank you to Zoltán Bagi, my second supervisor, who helped me in my daily work. He showed me how to use the instruments, how to carry out analyses and calculations and how to evaluate the results. His support was vital for my work.

I would like to express my thanks to Gábor Rákhely, the head of the Biotechnology Department who ensured my opportunity to work there during all this time. He was the one who offered me a PhD scholarship in the Environmental Sciences Doctoral school, wherein I stepped on the path to become a researcher. He as well, supported me in every regard and that he shared his insights into biotechnology with me was invaluable to my success.

I would like to thank Dénes Dudits, who envisaged the potential of willow as a biogas substrate and insisted the collaboration between the Dept. of Plant Biology (BRC, Szeged) and the Department of Biotechnology (University of Szeged). His great knowledge in biology and his friendly support helped me a great deal.

I owe a thank you to the department's other senior researchers who helped me with my work: Katalin Perei, whose advices regarding my wet-lab work were always very useful; and

Péter Kós, who, with his bioinformatics knowledge, helped me a great deal in when I first started to write code.

I would also like to thank my (ex-)colleagues who helped during my work: Márk Szuhaj, who contributed a great deal to the wet-lab part of the H_2-feeding experiment, Roland Wirth, who carried out the DNA and RNA isolation; and to Krisztián Laczi, Roland Tengölics, Norbert Ács and Árpád Szilágyi who also helped me in many ways.

I would like to express my thanks to other colleagues, laboratory technicians and fellow PhD workers, whom with at some point our careers paths crossed, helped me in my work or with their suggestions or friendly support and kind words (in no particular order): Ernő Homolya, Tamás Böjti, Etalka Kovács, Orsolya Strang, Szilvia Zsíros, Ágnes Kiss, Ágnes Dúzs, Sarolta Papp, Attila Bodor, Netta Bozóki, József Nyári, Barbara Hódi and Sarshad Koderi Valappil, and everyone else I may forgot to mention.

I also owe a big thank you to my current supervisor, Zsolt Boldogkői, who offered me a position in his research group at the Department of Medical Biology, and who supported me in finishing my thesis.

I would like to express my gratitude to my reviewers, Stefano Campanaro and Günther Bochmann, eminent scholars of the field, who devoted their time and energy to make this work better.

And last but certainly not the least, I would like to express my thanks, from the depths of my heart to my Family: my wonderful wife and my two amazing children. Years of working for a PhD takes a great toll not only on the candidate but on their loved ones as well. I'm so very lucky that they were always there for me, they supported me, and they showed patience when I had to work in the nights and in the weekends rather than spending time with them. I love you.

9. References

Ács, N., Bagi, Z., Rákhely, G., Minárovics, J., Nagy, K., Kovács, K.L., 2015. Bioaugmentation of biogas production by a hydrogen-producing bacterium. Bioresour. Technol. 186, 286–293. https://doi.org/10.1016/J.BIORTECH.2015.02.098

Ács, N., Szuhaj, M., Wirth, R., Bagi, Z., Maróti, G., Rákhely, G., Kovács, K.L., 2019. Microbial Community Rearrangements in Power-to-Biomethane Reactors Employing Mesophilic Biogas Digestate. Front. Energy Res. 7, 1–15. https://doi.org/10.3389/fenrg.2019.00132

Albertsen, M., Hugenholtz, P., Skarshewski, A., Nielsen, K.L., Tyson, G.W., Nielsen, P.H., 2013. Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. Nat. Biotechnol. 31, 533–538. https://doi.org/10.1038/nbt.2579

Alexander Martin, 1985. Biodegradation of organic chemicals. Environ. Sci. Technol. 19, 106–111.

Alexandropoulou, M., Antonopoulou, G., Fragkou, E., Ntaikou, I., Lyberatos, G., 2017. Fungal pretreatment of willow sawdust and its combination with alkaline treatment for enhancing biogas production. J. Environ. Manage. 203, 704–713. https://doi.org/10.1016/j.jenvman.2016.04.006

Allwright, M.R., Taylor, G., 2016. Molecular Breeding for Improved Second Generation Bioenergy Crops. Trends Plant Sci. 21, 43–54. https://doi.org/10.1016/j.tplants.2015.10.002

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. Nat. Methods 11, 1144–1146. https://doi.org/10.1038/nmeth.3103

Andrews, S., Krueger, F., Seconds-Pichon, A., Biggins, F., Wingett, S., 2010. FastQC. A quality control tool for high throughput sequence data [WWW Document]. Babraham

Inst. URL http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Asnicar, F., Thomas, A.M., Beghini, F., Mengoni, C., Manara, S., Manghi, P., Zhu, Q., Bolzan, M., Cumbo, F., May, U., Sanders, J.G., Zolfo, M., Kopylova, E., Pasolli, E., Knight, R., Mirarab, S., Huttenhower, C., Segata, N., 2020. Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. Nat. Commun. https://doi.org/10.1038/s41467-020-16366-7

AWG, 2019. The Anthropocene Working Group.

Bagi, Z., Ács, N., Böjti, T., Kakuk, B., Rákhely, G., Strang, O., Szuhaj, M., Wirth, R., Kovács, K.L.K.L., 2017. Biomethane: The energy storage, platform chemical and greenhouse gas mitigation target. Anaerobe 46, 13–22. https://doi.org/10.1016/j.anaerobe.2017.03.001

Bailera, M., Lisbona, P., Romeo, L.M., Espatolero, S., 2017. Power to Gas projects review: Lab, pilot and demo plants for storing renewable energy and CO2. Renew. Sustain. Energy Rev. https://doi.org/10.1016/j.rser.2016.11.130

Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O.U., Swartz, B., Quental, T.B., Marshall, C., McGuire, J.L., Lindsey, E.L., Maguire, K.C., Mersey, B., Ferrer, E.A., 2011. Has the Earth's sixth mass extinction already arrived? Nature. https://doi.org/10.1038/nature09678

Barua, S., Dhar, B.R., 2017. Advances towards understanding and engineering direct interspecies electron transfer in anaerobic digestion. Bioresour. Technol. 244, 698–707. https://doi.org/10.1016/j.biortech.2017.08.023

Batstone, D.J., Picioreanu, C., van Loosdrecht, M.C.M., 2006. Multidimensional modelling to investigate interspecies hydrogen transfer in anaerobic biofilms. Water Res. 40, 3099–3108. https://doi.org/10.1016/j.watres.2006.06.014

Böjti, T., Kovács, K.L.K.L.K.L., Kakuk, B., Wirth, R., Rákhely, G., Bagi, Z., 2017. Pretreatment of poultry manure for efficient biogas production as monosubstrate or co-fermentation with maize silage and corn stover. Anaerobe 46, 138–145. https://doi.org/10.1016/j.anaerobe.2017.03.017

Bonawitz, N.D., Chapple, C., 2013. Can genetic engineering of lignin deposition be accomplished without an unacceptable yield penalty? Curr. Opin. Biotechnol. 24, 336–343. https://doi.org/10.1016/j.copbio.2012.11.004

Borreani, G., Tabacco, E., Schmidt, R.J., Holmes, B.J., Muck, R.E., 2018. Silage review: Factors affecting dry matter and quality losses in silages. J. Dairy Sci. https://doi.org/10.3168/jds.2017-13837

Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T.B.K., Schulz, F., Jarett, J., Rivers, A.R., Eloe-Fadrosh, E.A., Tringe, S.G., Ivanova, N.N., Copeland, A., Clum, A., Becraft, E.D., Malmstrom, R.R., Birren, B., Podar, M., Bork, P., Weinstock, G.M., Garrity, G.M., Dodsworth, J.A., Yooseph, S., Sutton, G., Glöckner, F.O., Gilbert, J.A., Nelson, W.C., Hallam, S.J., Jungbluth, S.P., Ettema, T.J.G., Tighe, S., Konstantinidis, K.T., Liu, W.-T., Baker, B.J., Rattei, T., Eisen, J.A., Hedlund, B., McMahon, K.D., Fierer, N., Knight, R., Finn, R., Cochrane, G., Karsch-Mizrachi, I., Tyson, G.W., Rinke, C., Kyrpides, N.C., Schriml, L., Garrity, G.M., Hugenholtz, P., Sutton, G., Yilmaz, P., Meyer, F., Glöckner, F.O., Gilbert, J.A., Knight, R., Finn, R., Cochrane, G., Karsch-Mizrachi, I., Lapidus, A., Meyer, F., Yilmaz, P., Parks, D.H., Eren, A.M., Schriml, L., Banfield, J.F., Hugenholtz, P., Woyke, T., Woyke, T., 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat. Biotechnol. 35, 725–731. https://doi.org/10.1038/nbt.3893

Bremges, A., Maus, I., Belmann, P., Eikmeyer, F., Winkler, A., Albersmeier, A., Pühler, A., Schlüter, A., Sczyrba, A., 2015. Deeply sequenced metagenome and metatranscriptome of a biogas-producing microbial community from an agricultural production-scale biogas plant. Gigascience 4, 2–7. https://doi.org/10.1186/s13742-015-0073-6

Brethauer, S., Studer, M.H., 2015. Biochemical Conversion Processes of Lignocellulosic Biomass to Fuels and Chemicals – A Review. Chim. Int. J. Chem. https://doi.org/10.2533/chimia.2015.572

Brodeur, G., Yau, E., Badal, K., Collier, J., Ramachandran, K.B., Ramakrishnan, S., 2011. Chemical and Physicochemical Pretreatment of Lignocellulosic Biomass: A Review.

SAGE-Hindawi Access to Res. Enzym. Res. 787532. https://doi.org/10.4061/2011/787532

Cai, W., Liu, W., Wang, B., Yao, H., Guadie, A., Wang, A., 2020. Semiquantitative detection of hydrogen-associated or hydrogen-free electron transfer within methanogenic biofilm of microbial electrosynthesis. Appl. Environ. Microbiol. https://doi.org/10.1128/AEM.01056-20

Callahan, B.J., McMurdie, P.J., Holmes, S.P., 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME J. 11, 2639–2643. https://doi.org/10.1038/ismej.2017.119

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods 13, 581–583. https://doi.org/10.1038/nmeth.3869

Campanaro, S., Treu, L., Kougias, P.G., 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. https://doi.org/10.1186/s13068-016-0441-1

Campanaro, S., Treu, L., Kougias, P.G., 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. https://doi.org/10.1016/j.watres.2018.04.043

Campanaro, S., Treu, L., Rodriguez-R, L.M., Kovalovszki, A., Ziels, R.M., Maus, I., Zhu, X., Kougias, P.G., Basile, A., Luo, G., Schlüter, A., Konstantinidis, K.T., Angelidaki, I., 2020. New insights from the biogas microbiome by comprehensive genome-resolved metagenomics of nearly 1600 species originating from multiple anaerobic digesters. Biotechnol. Biofuels 2020 131 13, 1–18. https://doi.org/10.1186/S13068-020-01679-Y

Caputo, J., Balogh, S.B., Volk, T.A., Johnson, L., Puettmann, M., Lippke, B., Oneil, E., 2014. Incorporating Uncertainty into a Life Cycle Assessment (LCA) Model of Short-Rotation Willow Biomass (Salix spp.) Crops. BioEnergy Res. 7, 48–59.

https://doi.org/10.1007/s12155-013-9347-y

Ceballos, G., Ehrlich, P.R., Barnosky, A.D., García, A., Pringle, R.M., Palmer, T.M., 2015. Accelerated modern human-induced species losses: Entering the sixth mass extinction. Sci. Adv. https://doi.org/10.1126/sciadv.1400253

Chanoca, A., de Vries, L., Boerjan, W., 2019. Lignin Engineering in Forest Trees. Front. Plant Sci. 0, 912. https://doi.org/10.3389/FPLS.2019.00912

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, 1925–1927. https://doi.org/10.1093/bioinformatics/btz848

Cho, J.S., Kim, M.H., Bae, E.K., Choi, Y.I., Jeon, H.W., Han, K.H., Ko, J.H., 2021. Field evaluation of transgenic hybrid poplars with desirable wood properties and enhanced growth for biofuel production by bicistronic expression of PdGA20ox1 and PtrMYB3 in wood-forming tissue. Biotechnol. Biofuels. https://doi.org/10.1186/s13068-021-02029-2

Cho, S.-K., Im, W.-T., Kim, D.-H., Kim, M.-H., Shin, H.-S., Oh, S.-E., 2013. Dry anaerobic digestion of food waste under mesophilic conditions: Performance and methanogenic community analysis. Bioresour. Technol. 131, 210–217. https://doi.org/10.1016/J.BIORTECH.2012.12.100

Clifton-Brown, J., Harfouche, A., Casler, M.D., Dylan Jones, H., Macalpine, W.J., Murphy-Bokern, D., Smart, L.B., Adler, A., Ashman, C., Awty-Carroll, D., Bastien, C., Bopper, S., Botnari, V., Brancourt-Hulmel, M., Chen, Z., Clark, L. V., Cosentino, S., Dalton, S., Davey, C., Dolstra, O., Donnison, I., Flavell, R., Greef, J., Hanley, S., Hastings, A., Hertzberg, M., Hsu, T.-W., Huang, L.S., Iurato, A., Jensen, E., Jin, X., Jørgensen, U., Kiesel, A., Kim, D.-S., Liu, J., McCalmont, J.P., McMahon, B.G., Mos, M., Robson, P., Sacks, E.J., Sandu, A., Scalici, G., Schwarz, K., Scordia, D., Shafiei, R., Shield, I., Slavov, G., Stanton, B.J., Swaminathan, K., Taylor, G., Torres, A.F., Trindade, L.M., Tschaplinski, T., Tuskan, G.A., Yamada, T., Yeon Yu, C., Zalesny, R.S., Zong, J., Lewandowski, I., 2019. Breeding progress and preparedness for mass-scale deployment of perennial lignocellulosic biomass crops switchgrass, miscanthus, willow

and poplar. GCB Bioenergy 11, 118–151. https://doi.org/10.1111/gcbb.12566

Cozzi, L., Goodson, T., 2020. Empowering electricity consumers to lower their carbon footprint.

Cunniff, J., Cerasuolo, M., 2011. Lighting the way to willow biomass production. J. Sci. Food Agric. 91, 1733–1736. https://doi.org/10.1002/jsfa.4470

Dimitriou, I., Mola-Yudego, B., Aronsson, P., Eriksson, J., 2012. Changes in Organic Carbon and Trace Elements in the Soil of Willow Short-Rotation Coppice Plantations. BioEnergy Res. 5, 563–572. https://doi.org/10.1007/s12155-012-9215-1

Dudits, D., Cseri, A., Török, K., Sass, L., Zombori, Z., Ferenc, G., Poór, P., Borbély, P., Czékus, Z., Vankova, R., Dobrev, P., Szántó, J., Bagi, Z., Kovács, K.L., 2022. Triploid Hybrid Vigor in Above-Ground Growth and Methane Fermentation Efficiency of Energy Willow. Front. Plant Sci. 13, 1–14. https://doi.org/10.3389/fpls.2022.770284

Dudits, D., Török, K., Cseri, A., Paul, K., Nagy, A.V., Nagy, B., Sass, L., Ferenc, G., Vankova, R., Dobrev, P., Vass, I., Ayaydin, F., 2016. Response of Organ Structure and Physiology to Autotetraploidization in Early Development of Energy Willow Salix viminalis, Plant Physiology. https://doi.org/10.1104/pp.15.01679

Eisenbies, M.H., Volk, T.A., Posselius, J., Foster, C., Shi, S., Karapetyan, S., 2014. Evaluation of a Single-Pass, Cut and Chip Harvest System on Commercial-Scale, Short-Rotation Shrub Willow Biomass Crops. BioEnergy Res. 7, 1506–1518. https://doi.org/10.1007/s12155-014-9482-0

Eren, A.M., Esen, Ö.C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., Delmont, T.O., 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ 3, e1319. https://doi.org/10.7717/peerj.1319

Estevez, M.M., Linjordet, R., Morken, J., 2012. Effects of steam explosion and co-digestion in the methane production from Salix by mesophilic batch assays. Bioresour. Technol. 104, 749–756. https://doi.org/10.1016/j.biortech.2011.11.017

Fabio, E.S., Volk, T.A., Miller, R.O., Serapiglia, M.J., Gauch, H.G., Van Rees, K.C.J., Hangs, R.D., Amichev, B.Y., Kuzovkina, Y.A., Labrecque, M., Johnson, G.A., Ewy,

R.G., Kling, G.J., Smart, L.B., 2017. Genotype × environment interaction analysis of North American shrub willow yield trials confirms superior performance of triploid hybrids. GCB Bioenergy 9, 445–459. https://doi.org/10.1111/gcbb.12344

Ferrarini, A., Fornasier, F., Serra, P., Ferrari, F., Trevisan, M., Amaducci, S., 2017. Impacts of willow and miscanthus bioenergy buffers on biogeochemical N removal processes along the soil-groundwater continuum. GCB Bioenergy 9, 246–261. https://doi.org/10.1111/gcbb.12340

Fischer, G., Prieler, S., van Velthuizen, H., 2005. Biomass potentials of miscanthus, willow and poplar: results and policy implications for Eastern Europe, Northern and Central Asia. Biomass and Bioenergy 28, 119–132. https://doi.org/10.1016/J.BIOMBIOE.2004.08.013

Folke, C., Polasky, S., Rockström, J., Galaz, V., Westley, F., Lamont, M., Scheffer, M., Österblom, H., Carpenter, S.R., Chapin, F.S., Seto, K.C., Weber, E.U., Crona, B.I., Daily, G.C., Dasgupta, P., Gaffney, O., Gordon, L.J., Hoff, H., Levin, S.A., Lubchenco, J., Steffen, W., Walker, B.H., 2021. Our future in the Anthropocene biosphere. Ambio. https://doi.org/10.1007/s13280-021-01544-8

Fontana, A., Kougias, P.G., Treu, L., Kovalovszki, A., Valle, G., Cappa, F., Morelli, L., Angelidaki, I., Campanaro, S., 2018. Microbial activity response to hydrogen injection in thermophilic anaerobic digesters revealed by genome-centric metatranscriptomics. Microbiome 6, 194. https://doi.org/10.1186/s40168-018-0583-4

Frei, M., 2013. Lignin: characterization of a multifaceted crop component. ScientificWorldJournal. 2013, 436517. https://doi.org/10.1155/2013/436517

Frigon, J.-C., Mehta, P., Guiot, S.R., 2012. Impact of mechanical, chemical and enzymatic pre-treatments on the methane yield from the anaerobic digestion of switchgrass. Biomass and Bioenergy 36, 1–11. https://doi.org/10.1016/j.biombioe.2011.02.013

Fuksa, P., Hakl, J., Míchal, P., Hrevušová, Z., Šantrůček, J., Tlustoš, P., 2020. Effect of silage maize plant density and plant parts on biogas production and composition. Biomass and Bioenergy 142, 105770. https://doi.org/10.1016/j.biombioe.2020.105770

Gissén, C., Prade, T., Kreuger, E., Nges, I.A., Rosenqvist, H., Svensson, S.-E.E., Lantz, M., Mattsson, J.E., Börjesson, P., Björnsson, L., 2014. Comparing energy crops for biogas production – Yields, energy input and costs in cultivation using digestate and mineral fertilisation. Biomass and Bioenergy 64, 199–210. https://doi.org/10.1016/j.biombioe.2014.03.061

Goering, H.K., Van Soest, P.J., 1970. Forage fiber analyses.

Graham, E.D., Heidelberg, J.F., Tully, B.J., 2017. BinSanity: unsupervised clustering of environmental microbial assemblies using coverage and affinity propagation. PeerJ 5, e3035. https://doi.org/10.7717/peerj.3035

Grochowski, L.L., White, R.H., 2010. Biosynthesis of the methanogenic coenzymes. Compr. Nat. Prod. II Chem. Biol. 7, 711–748. https://doi.org/10.1016/b978-008045382-8.00147-7

Güllert, S., Fischer, M.A., Turaev, D., Noebauer, B., Ilmberger, N., Wemheuer, B., Alawi, M., Rattei, T., Daniel, R., Schmitz, R.A., Grundhoff, A., Streit, W.R., 2016. Deep metagenome and metatranscriptome analyses of microbial communities affiliated with an industrial biogas fermenter, a cow rumen, and elephant feces reveal major differences in carbohydrate hydrolysis strategies. Biotechnol. Biofuels. https://doi.org/10.1186/s13068-016-0534-x

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

Hashimoto, K., 1994. Metastable metals for “green” materials for global atmosphere conservation and abundant energy supply. Mater. Sci. Eng. A. https://doi.org/10.1016/0921-5093(94)90158-9

Hashimoto, K., Yamasaki, M., Fujimura, K., Matsui, T., Izumiya, K., Komori, M., El-Moneim, A.A., Akiyama, E., Habazaki, H., Kumagai, N., Kawashima, A., Asami, K., 1999. Global CO2 recycling - Novel materials and prospect for prevention of global warming and abundant energy supply. Mater. Sci. Eng. A.

https://doi.org/10.1016/S0921-5093(99)00092-1

Hassa, J., Maus, I., Off, S., Pühler, A., Scherer, P., Klocke, M., Schlüter, A., 2018. Metagenome, metatranscriptome, and metaproteome approaches unraveled compositions and functional relationships of microbial communities residing in biogas plants. Appl. Microbiol. Biotechnol. 102, 5045–5063. https://doi.org/10.1007/s00253-018-8976-7

Hendriks, a. T.W.M., Zeeman, G., 2009. Pretreatments to enhance the digestibility of lignocellulosic biomass. Bioresour. Technol. 100, 10–8. https://doi.org/10.1016/j.biortech.2008.05.027

Herrmann, C., Idler, C., Heiermann, M., 2016. Biogas crops grown in energy crop rotations: Linking chemical composition and methane production characteristics. Bioresour. Technol. 206, 23–35. https://doi.org/10.1016/j.biortech.2016.01.058

Horn, S.J., Estevez, M.M., Nielsen, H.K., Linjordet, R., Eijsink, V.G.H., 2011. Biogas production and saccharification of Salix pretreated at different steam explosion conditions. Bioresour. Technol. 102, 7932–7936. https://doi.org/10.1016/j.biortech.2011.06.042

Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., von Mering, C., Bork, P., 2017. Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. Mol. Biol. Evol. 34, 2115–2122. https://doi.org/10.1093/molbev/msx148

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., Mende, D.R., Letunic, I., Rattei, T., Jensen, L.J., von Mering, C., Bork, P., 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. 47, D309–D314. https://doi.org/10.1093/nar/gky1085

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. https://doi.org/10.1186/1471-2105-11-119

IEA, 2021. Global Energy Review 2021 – Analysis - IEA. Int. Energy Agency.

Isroi, Millati, R., Syamsiah, S., Niklasson, C., Cahyanto, M.N., Lundquist, K., Taherzadeh, M.J., Isroi, ., Millati, R., Syamsiah, S., Niklasson, C., Cahyanto, M.N., Ludquist, K., Taherzadeh, M.J., 2011. Biological pretreatment of lignocelluloses with white-rot fungi and its applications: A review, BioResources. https://doi.org/10.15376/biores.6.4.5224-5259

Jaenicke, S., Ander, C., Bekel, T., Bisdorf, R., Dröge, M., Gartemann, K.-H.H., Jünemann, S., Kaiser, O., Krause, L., Tille, F., Zakrzewski, M., Pühler, A., Schlüter, A., Goesmann, A., 2011. Comparative and Joint Analysis of Two Metagenomic Datasets from a Biogas Fermenter Obtained by 454-Pyrosequencing. PLoS One 6, e14519. https://doi.org/10.1371/journal.pone.0014519

Jankowski, K.J., Dubis, B., Sokólski, M.M., Załuski, D., Bórawski, P., Szempliński, W., 2020. Productivity and energy balance of maize and sorghum grown for biogas in a large-area farm in Poland: An 11-year field experiment. Ind. Crops Prod. 148, 112326. https://doi.org/10.1016/j.indcrop.2020.112326

Janssen, J., Weyens, N., Croes, S., Beckers, B., Meiresonne, L., Van Peteghem, P., Carleer, R., Vangronsveld, J., 2015. Phytoremediation of Metal Contaminated Soil Using Willow: Exploiting Plant-Associated Bacteria to Improve Biomass Production and Metal Uptake. Int. J. Phytoremediation 17, 1123–1136. https://doi.org/10.1080/15226514.2015.1045129

Jia, Y., Ng, S.-K., Lu, H., Cai, M., Lee, P.K.H., 2018. Genome-centric metatranscriptomes and ecological roles of the active microbial populations during cellulosic biomass anaerobic digestion. Biotechnol. Biofuels 11, 117. https://doi.org/10.1186/s13068-018-1121-0

Johnson, J.S., Spakowicz, D.J., Hong, B.-Y., Petersen, L.M., Demkowicz, P., Chen, L., Leopold, S.R., Hanson, B.M., Agresta, H.O., Gerstein, M., Sodergren, E., Weinstock, G.M., 2019. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nat. Commun. 2019 101 10, 1–11. https://doi.org/10.1038/s41467-019-13036-1

Jonsson, V., Österlund, T., Nerman, O., Kristiansson, E., 2016. Statistical evaluation of methods for identification of differentially abundant genes in comparative metagenomics. BMC Genomics 17, 78. https://doi.org/10.1186/s12864-016-2386-y

Jünemann, S., Kleinbölting, N., Jaenicke, S., Henke, C., Hassa, J., Nelkner, J., Stolze, Y., Albaum, S.P., Schlüter, A., Goesmann, A., Sczyrba, A., Stoye, J., 2017. Bioinformatics for NGS-based metagenomics and the application to biogas research. J. Biotechnol. 261, 10–23. https://doi.org/10.1016/J.JBIOTEC.2017.08.012

Jurado, E., Gavala, H.N., Skiadas, I. V., 2013. Enhancement of methane yield from wheat straw, miscanthus and willow using aqueous ammonia soaking. Environ. Technol. (United Kingdom) 34, 2069–2075. https://doi.org/10.1080/09593330.2013.826701

Kakuk, B., Bagi, Z., Rákhely, G., Maróti, G., Dudits, D., Kovács, K.L., 2021a. Methane production from green and woody biomass using short rotation willow genotypes for bioenergy generation. Bioresour. Technol. 333. https://doi.org/10.1016/j.biortech.2021.125223

Kakuk, B., Kovács, K.L., Szuhaj, M., Rákhely, G., Bagi, Z., 2017. Adaptation of continuous biogas reactors operating under wet fermentation conditions to dry conditions with corn stover as substrate. Anaerobe 46, 78–85. https://doi.org/10.1016/j.anaerobe.2017.05.015

Kakuk, B., Wirth, R., Maróti, G., Szuhaj, M., Rákhely, G., Laczi, K., Kovács, K.L., Bagi, Z., 2021b. Early response of methanogenic archaea to H2 as evaluated by metagenomics and metatranscriptomics. Microb. Cell Fact. 20. https://doi.org/10.1186/s12934-021-01618-y

Kalamaras, S.D., Kotsopoulos, T.A., 2014. Anaerobic co-digestion of cattle manure and alternative crops for the substitution of maize in South Europe. Bioresour. Technol. 172. https://doi.org/10.1016/j.biortech.2014.09.005

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., Tanabe, M., 2016. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. https://doi.org/10.1093/nar/gkv1070

Kang, D.D., Froula, J., Egan, R., Wang, Z., 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ 3, e1165. https://doi.org/10.7717/peerj.1165

Kassambara, A., Mundt, F., Kassambara, A.; Mundt, F., 2017. Factoextra: extract and visualize the results of multivariate data analyses. URL http//www.sthda.com/english/rpkgs/factoextra BugReports.

Khlystov, N.A., Yoshikuni, Y., Deutsch, S., Sattely, E.S., 2021. A plant host, Nicotiana benthamiana, enables the production and study of fungal lignin-degrading enzymes. Commun. Biol. 2021 41 4, 1–13. https://doi.org/10.1038/s42003-021-02464-9

Kiesel, A., Nunn, C., Iqbal, Y., Van der Weijde, T., Wagner, M., Özgüven, M., Tarakanov, I., Kalinina, O., Trindade, L.M., Clifton-Brown, J., Lewandowski, I., 2017. Site-Specific Management of Miscanthus Genotypes for Combustion and Anaerobic Digestion: A Comparison of Energy Yields. Front. Plant Sci. 8, 347. https://doi.org/10.3389/fpls.2017.00347

Koeck, D.E., Hahnke, S., Zverlov, V. V, 2016. Herbinix luporum sp. nov., a thermophilic cellulose-degrading bacterium isolated from a thermophilic biogas reactor. Int. J. Syst. Evol. Microbiol. 66, 4132–4137. https://doi.org/10.1099/ijsem.0.001324

Konstantinidis, K.T., Tiedje, J.M., 2005. Towards a Genome-Based Taxonomy for Prokaryotes. J. Bacteriol. 187, 6258–6264. https://doi.org/10.1128/JB.187.18.6258-6264.2005

Kotsyurbenko, O.R., Glagolev, M. V., Nozhevnikova, A.N., Conrad, R., 2001. Competition between homoacetogenic bacteria and methanogenic archaea for hydrogen at low temperature. FEMS Microbiol. Ecol. 38, 153–159. https://doi.org/10.1111/j.1574-6941.2001.tb00893.x

Kougias, P.G., Angelidaki, I., 2018. Biogas and its opportunities—A review. Front. Environ. Sci. Eng. 12. https://doi.org/10.1007/s11783-018-1037-8

Kougias, P.G., Campanaro, S., Treu, L., Tsapekos, P., Armani, A., Angelidaki, I., 2018. Spatial Distribution and Diverse Metabolic Functions of Lignocellulose-Degrading

Uncultured Bacteria as Revealed by Genome-Centric Metagenomics. Appl. Environ. Microbiol. 84, AEM.01244-18. https://doi.org/10.1128/AEM.01244-18

Kovács, K.L., Ács, N., Kovács, E., Wirth, R., Rákhely, G., Strang, O., Herbel, Z., Bagi, Z., 2013. Improvement of biogas production by bioaugmentation. Biomed Res. Int. 2013, 482653. https://doi.org/10.1155/2013/482653

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923

Lê, S., Josse, J., Husson, F., 2008. FactoMineR: An R package for multivariate analysis. J. Stat. Softw. https://doi.org/10.18637/jss.v025.i01

Li, A., Chu, Y., Wang, X., Ren, L., Yu, J., Liu, X., Yan, J., Zhang, L., Wu, S., Li, S., 2013. A pyrosequencing-based metagenomic study of methane-producing microbial community in solid-state biogas reactor. Biotechnol. Biofuels 6, 3. https://doi.org/10.1186/1754-6834-6-3

Li, D., Luo, R., Liu, C.-M., Leung, C.-M., Ting, H.-F., Sadakane, K., Yamashita, H., Lam, T.-W., 2016. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods 102, 3–11. https://doi.org/10.1016/J.YMETH.2016.02.020

Li, L., He, Q., Ma, Y., Wang, X., Peng, X., 2016. A mesophilic anaerobic digester for treating food waste: Process stability and microbial community analysis using pyrosequencing. Microb. Cell Fact. 15, 1–11. https://doi.org/10.1186/s12934-016-0466-y

Li, Y.F., Nelson, M.C., Chen, P.H., Graf, J., Li, Y., Yu, Z., 2014. Comparison of the microbial communities in solid-state anaerobic digestion (SS-AD) reactors operated at mesophilic and thermophilic temperatures. Appl. Microbiol. Biotechnol. https://doi.org/10.1007/s00253-014-6036-5

Liao, Y., Smyth, G.K., Shi, W., 2019. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res. https://doi.org/10.1093/nar/gkz114

Liebetrau, J., Baier, U., Wall, D., Murphy, J.D., 2020. Integration of biogas systems into the energy system. IEA Bioenergy.

Ligaba-Osena, A., Hankoua, B., DiMarco, K., Pace, R., Crocker, M., McAtee, J., Nagachar, N., Tien, M., Richard, T.L., 2017. Reducing biomass recalcitrance by heterologous expression of a bacterial peroxidase in tobacco (Nicotiana benthamiana). Sci. Reports 2017 71 7, 1–18. https://doi.org/10.1038/s41598-017-16909-x

Lin, Q., De Vrieze, J., He, G., Li, X., Li, J., 2016. Temperature regulates methane production through the function centralization of microbial community in anaerobic digestion. Bioresour. Technol. 216, 150–158. https://doi.org/10.1016/j.biortech.2016.05.046

Littlewood, J., Guo, M., Boerjan, W., Murphy, R.J., 2014. Bioethanol from poplar: a commercially viable alternative to fossil fuel in the European Union. Biotechnol. Biofuels 7. https://doi.org/10.1186/1754-6834-7-113

Logroño, W., Popp, D., Nikolausz, M., Kluge, P., Harms, H., Kleinsteuber, S., 2021. Microbial Communities in Flexible Biomethanation of Hydrogen Are Functionally Resilient Upon Starvation. Front. Microbiol. 12, 1–12. https://doi.org/10.3389/fmicb.2021.619632

López, M.J., Suárez-Estrella, F., Vargas-García, M.C., López-González, J.A., Verstichel, S., Debeer, L., Wierinck, I., Moreno, J., 2013. Biodelignification of agricultural and forest wastes: Effect on anaerobic digestion. Biomass and Bioenergy 58, 343–349. https://doi.org/10.1016/j.biombioe.2013.10.021

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Maus, I., Koeck, D.E., Cibis, K.G., Hahnke, S., Kim, Y.S., Langer, T., Kreubel, J., Erhard, M., Bremges, A., Off, S., Stolze, Y., Jaenicke, S., Goesmann, A., Sczyrba, A., Scherer, P., König, H., Schwarz, W.H., Zverlov, V. V., Liebl, W., Pühler, A., Schlüter, A., Klocke, M., 2016. Unraveling the microbiome of a thermophilic biogas plant by metagenome and metatranscriptome analysis complemented by characterization of

bacterial and archaeal isolates. Biotechnol. Biofuels 9, 171. https://doi.org/10.1186/s13068-016-0581-3

Maus, I., Wibberg, D., Stantscheff, R., Stolze, Y., Blom, J., Eikmeyer, F.G., Fracowiak, J., König, H., Pühler, A., Schlüter, A., 2015. Insights into the annotated genome sequence of Methanoculleus bourgensis MS2T, related to dominant methanogens in biogas-producing plants. J. Biotechnol. 201, 43–53. https://doi.org/10.1016/j.jbiotec.2014.11.020

McGhee, T., 1968. A method for approximation of the volatile acid concentrations in anaerobic digesters. Water Sew. Work. 115, 162–166.

Montgomery, L.F.R., Bochmann, G., 2014. Pretreatment of feedstock for enhanced biogas production. IEA Bioenergy Ireland.

Mulat, D.G., Dibdiakova, J., Horn, S.J., 2018a. Microbial biogas production from hydrolysis lignin: insight into lignin structural changes. Biotechnol. Biofuels 11, 61. https://doi.org/10.1186/s13068-018-1054-7

Mulat, D.G., Huerta, S.G., Kalyani, D., Horn, S.J., 2018b. Enhancing methane production from lignocellulosic biomass by combined steam-explosion pretreatment and bioaugmentation with cellulolytic bacterium Caldicellulosiruptor bescii. Biotechnol. Biofuels 11, 19. https://doi.org/10.1186/s13068-018-1025-z

Müller-Langer, F., Majer, S., O'Keeffe, S., 2014. Benchmarking biofuels—a comparison of technical, economic and environmental indicators. Energy. Sustain. Soc. 4, 1–14. https://doi.org/10.1186/s13705-014-0020-x

Narayanasamy, S., Jarosz, Y., Muller, E.E.L., Heintz-Buschart, A., Herold, M., Kaysen, A., Laczny, C.C., Pinel, N., May, P., Wilmes, P., 2016. IMP: a pipeline for reproducible reference-independent integrated metagenomic and metatranscriptomic analyses. Genome Biol. 17, 260. https://doi.org/10.1186/s13059-016-1116-8

Nielsen, H.B., Almeida, M., Juncker, A.S., Rasmussen, S., Li, J., Sunagawa, S., Plichta, D.R., Gautier, L., Pedersen, A.G., Le Chatelier, E., Pelletier, E., Bonde, I., Nielsen, T., Manichanh, C., Arumugam, M., Batto, J.M., Quintanilha Dos Santos, M.B., Blom,

N., Borruel, N., Burgdorf, K.S., Boumezbeur, F., Casellas, F., Doré, J., Dworzynski, P., Guarner, F., Hansen, T., Hildebrand, F., Kaas, R.S., Kennedy, S., Kristiansen, K., Kultima, J.R., Léonard, P., Levenez, F., Lund, O., Moumen, B., Le Paslier, D., Pons, N., Pedersen, O., Prifti, E., Qin, J., Raes, J., Sørensen, S., Tap, J., Tims, S., Ussery, D.W., Yamada, T., Renault, P., Sicheritz-Ponten, T., Bork, P., Wang, J., Brunak, S., Ehrlich, S.D., 2014. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. Nat. Biotechnol. 32, 822–828. https://doi.org/10.1038/nbt.2939

Nobu, M.K., Narihiro, T., Rinke, C., Kamagata, Y., Tringe, S.G., Woyke, T., Liu, W.-T., 2015. Microbial dark matter ecogenomics reveals complex synergistic networks in a methanogenic bioreactor. ISME J. 9, 1710–1722. https://doi.org/10.1038/ismej.2014.256

Nurk, L., Bühle, L., Wachendorf, M., 2016. Degradation of Fibre and Non-fibre Fractions During Anaerobic Digestion in Silages of Maize, Sunflower and Sorghum-Sudangrass of Different Maturities. BioEnergy Res. 9, 720–730. https://doi.org/10.1007/s12155-016-9717-3

Nyári, J., Kakuk, B., Bagi, Z., Rákhely, G., Kovács, K.L., 2021. Use of ensiled green willow biomass in biogas fermentation. Biol. Futur. 72, 263–271. https://doi.org/10.1007/s42977-021-00067-3

Parker, C.H., Keefe, E.R., Herzog, N.M., O'connell, J.F., Hawkes, K., 2016. The pyrophilic primate hypothesis. Evol. Anthropol. https://doi.org/10.1002/evan.21475

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 Res. 25, 1043–55. https://doi.org/10.1101/gr.186072.114

Patinvoh, R.J., Osadolor, O.A., Chandolias, K., Sárvári Horváth, I., Taherzadeh, M.J., 2017. Innovative pretreatment strategies for biogas production. Bioresour. Technol. https://doi.org/10.1016/j.biortech.2016.11.083

Petzoldt, T., 2018. growthrates: Estimate Growth Rates from Experimental Data.

Pollock, J., Glendinning, L., Wisedchanwet, T., Watson, M., 2018. The Madness of Microbiome: Attempting To Find Consensus "Best Practice" for 16S Microbiome Studies. Appl. Environ. Microbiol. 84, e02627-17. https://doi.org/10.1128/AEM.02627-17

Prodan, A., Tremaroli, V., Brolin, H., Zwinderman, A.H., Nieuwdorp, M., Levin, E., 2020. Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. PLoS One 15, e0227434. https://doi.org/10.1371/JOURNAL.PONE.0227434

Purdy, S.J., Maddison, A.L., Nunn, C.P., Winters, A., Timms-Taravella, E., Jones, C.M., Clifton-Brown, J.C., Donnison, I.S., Gallagher, J.A., 2017. Could Miscanthus replace maize as the preferred substrate for anaerobic digestion in the United Kingdom? Future breeding strategies. Glob. Change Biol. Bioenergy 9, 1122–1139. https://doi.org/10.1111/gcbb.12419

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41, D590-6. https://doi.org/10.1093/nar/gks1219

R Core Team, 2019. R: A Language and Environment for Statistical Computing.

Rafrafi, Y., Laguillaumie, L., Dumas, C., 2020. Biological Methanation of H2 and CO2 with Mixed Cultures: Current Advances, Hurdles and Challenges. Waste and Biomass Valorization. https://doi.org/10.1007/s12649-020-01283-z

Ray, M.J., Brereton, N.J.B., Shield, I., Karp, A., Murphy, R.J., 2012. Variation in Cell Wall Composition and Accessibility in Relation to Biofuel Potential of Short Rotation Coppice Willows. Bioenergy Res. 5, 685–698. https://doi.org/10.1007/s12155-011-9177-8

Rodriguez-R, L.M., Gunturu, S., Harvey, W.T., Rosselló-Mora, R., Tiedje, J.M., Cole, J.R., Konstantinidis, K.T., 2018. The Microbial Genomes Atlas (MiGA) webserver: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. Nucleic Acids Res. 46, W282–W288. https://doi.org/10.1093/nar/gky467

Rodriguez-R, L.M., Konstantinidis, K.T., 2014. Bypassing Cultivation To Identify Bacterial Species. Microbe Mag. 9, 111–118. https://doi.org/10.1128/microbe.9.111.1

Rotaru, A.E., Shrestha, P.M., Liu, F., Shrestha, M., Shrestha, D., Embree, M., Zengler, K., Wardman, C., Nevin, K.P., Lovley, D.R., 2014. A new model for electron flow during anaerobic digestion: Direct interspecies electron transfer to Methanosaeta for the reduction of carbon dioxide to methane. Energy Environ. Sci. 7, 408–415. https://doi.org/10.1039/c3ee42189a

Rouches, E., Herpoël-Gimbert, I., Steyer, J.P., Carrere, H., 2016a. Improvement of anaerobic degradation by white-rot fungi pretreatment of lignocellulosic biomass: A review. Renew. Sustain. Energy Rev. 59, 179–198. https://doi.org/10.1016/J.RSER.2015.12.317

Rouches, E., Zhou, S., Steyer, J.P., Carrere, H., 2016b. White-Rot Fungi pretreatment of lignocellulosic biomass for anaerobic digestion: Impact of glucose supplementation. Process Biochem. 51, 1784–1792.

Scarlat, N., Dallemand, J.F., Fahl, F., 2018. Biogas: Developments and perspectives in Europe. Renew. Energy. https://doi.org/10.1016/j.renene.2018.03.006

Schlüter, A., Bekel, T., Diaz, N.N., Dondrup, M., Eichenlaub, R., Gartemann, K.-H.H., Krahn, I., Krause, L., Krömeke, H., Kruse, O., Mussgnug, J.H., Neuweger, H., Niehaus, K., Pühler, A., Runte, K.J., Szczepanowski, R., Tauch, A., Tilker, A., Viehöver, P., Goesmann, A., 2008. The metagenome of a biogas-producing microbial community of a production-scale biogas plant fermenter analysed by the 454-pyrosequencing technology. J. Biotechnol. 136, 77–90. https://doi.org/10.1016/j.jbiotec.2008.05.008

Schroyen, M., Vervaeren, H., Van Hulle, S.W.H., Raes, K., 2014. Impact of enzymatic pretreatment on corn stover degradation and biogas production. Bioresour. Technol. 173C, 59–66. https://doi.org/10.1016/j.biortech.2014.09.030

Schroyen, M., Vervaeren, H., Vandepitte, H., Van Hulle, S.W.H., Raes, K., 2015. Effect of enzymatic pretreatment of various lignocellulosic substrates on production of phenolic

compounds and biomethane potential. Bioresour. Technol. 192. https://doi.org/10.1016/j.biortech.2015.06.051

Secretariat REN21, 2021. RENEWABLES 2021 GLOBAL STATUS REPORT 2021, Global Status Report for Buildings and Construction: Towards a Zero-emission, Efficient and Resilient Buildings and Construction Sector.

Serapiglia, M.J., Gouker, F.E., Smart, L.B., 2014. Early selection of novel triploid hybrids of shrub willow with improved biomass yield relative to diploids. BMC Plant Biol. 14, 74. https://doi.org/10.1186/1471-2229-14-74

Serapiglia, M.J., Humiston, M.C., Xu, H., Hogsett, D.A., de Orduña, R.M., Stipanovic, A.J., Smart, L.B., Orduña, R.M. de, Stipanovic, A.J., Smart, L.B., 2013. Enzymatic saccharification of shrub willow genotypes with differing biomass composition for biofuel production 4, 57. https://doi.org/10.3389/fpls.2013.00057

Shah, F.A., Mahmood, Q., Rashid, N., Pervez, A., Raja, I.A., Shah, M.M., 2015. Co-digestion, pretreatment and digester design for enhanced methanogenesis. Renew. Sustain. Energy Rev. 42, 627–642. https://doi.org/10.1016/j.rser.2014.10.053

Shakya, M., Lo, C.-C., Chain, P.S.G., 2019. Advances and Challenges in Metatranscriptomic Analysis. Front. Genet. 10, 904. https://doi.org/10.3389/FGENE.2019.00904

Shimizu, S., Upadhye, R., Ishijima, Y., Naganuma, T., 2011. Methanosarcina horonobensis sp. nov., a methanogenic archaeon isolated from a deep subsurface miocene formation. Int. J. Syst. Evol. Microbiol. 61, 2503–2507. https://doi.org/10.1099/ijs.0.028548-0

Sieber, C.M.K., Probst, A.J., Sharrar, A., Thomas, B.C., Hess, M., Tringe, S.G., Banfield, J.F., 2018. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. Nat. Microbiol. 3, 836–843. https://doi.org/10.1038/s41564-018-0171-1

Simon, P., Deák, P., 2020. RENEWABLE ENERGY LAW AND REGULATION IN HUNGARY.

Sleight, N.J., Volk, T.A., 2016. Recently Bred Willow (Salix spp.) Biomass Crops Show Stable Yield Trends Over Three Rotations at Two Sites. Bioenergy Res. 9, 782–797. https://doi.org/10.1007/s12155-016-9726-2

Ssegane, H., Zumpf, C., Cristina Negri, M., Campbell, P., Heavey, J.P., Volk, T.A., 2016. The economics of growing shrub willow as a bioenergy buffer on agricultural fields: A case study in the Midwest Corn Belt. Biofuels, Bioprod. Biorefining 10, 776–789. https://doi.org/10.1002/bbb.1679

Steffen, W., Rockström, J., Richardson, K., Lenton, T.M., Folke, C., Liverman, D., Summerhayes, C.P., Barnosky, A.D., Cornell, S.E., Crucifix, M., Donges, J.F., Fetzer, I., Lade, S.J., Scheffer, M., Winkelmann, R., Schellnhuber, H.J., 2018. Trajectories of the Earth System in the Anthropocene. Proc. Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.1810141115

Stolarski, M.J., Niksa, D., Krzyżaniak, M., Tworkowski, J., Szczukowski, S., 2019. Willow productivity from small- and large-scale experimental plantations in Poland from 2000 to 2017. Renew. Sustain. Energy Rev. 101, 461–475. https://doi.org/10.1016/J.RSER.2018.11.034

Stolarski, M.J., Szczukowski, S., Tworkowski, J., Krzyżaniak, M., Załuski, D., Ozef Tworkowski, J., Krzy, M., Zaniak, _, Załuski, D., 2017. Willow biomass and cuttings' production potential over ten successive annual harvests. Biomass and Bioenergy 105, 230–247. https://doi.org/10.1016/j.biombioe.2017.07.010

Stolze, Y., Bremges, A., Rumming, M., Henke, C., Maus, I., Pühler, A., Sczyrba, A., Schlüter, A., 2016. Identification and genome reconstruction of abundant distinct taxa in microbiomes from one thermophilic and three mesophilic production-scale biogas plants. Biotechnol. Biofuels 9, 156. https://doi.org/10.1186/s13068-016-0565-3

Stolze, Y., Zakrzewski, M., Maus, I., Eikmeyer, F., Jaenicke, S., Rottmann, N., Siebner, C., Pühler, A., Schlüter, A., 2015. Comparative metagenomics of biogas-producing microbial communities from production-scale biogas plants operating under wet or dry fermentation conditions. Biotechnol. Biofuels 8, 14. https://doi.org/10.1186/s13068-014-0193-8

Stoof, C.R., Richards, B.K., Woodbury, P.B., Fabio, E.S., Brumbach, A.R., Cherney, J., Das, S., Geohring, L., Hansen, J., Hornesky, J., Mayton, H., Mason, C., Ruestow, G., Smart, L.B., Volk, T.A., Steenhuis, T.S., 2015. Untapped Potential: Opportunities and Challenges for Sustainable Bioenergy Production from Marginal Lands in the Northeast USA. BioEnergy Res. 8, 482–501. https://doi.org/10.1007/s12155-014-9515-8

Sun, Q., Li, H., Yan, J., Liu, L., Yu, Z., Yu, X., 2015. Selection of appropriate biogas upgrading technology-a review of biogas cleaning, upgrading and utilisation. Renew. Sustain. Energy Rev. 51, 521–532. https://doi.org/10.1016/j.rser.2015.06.029

Szuhaj, M., Ács, N., Tengölics, R., Bodor, A., Rákhely, G., Kovács, K.L., Bagi, Z., 2016. Conversion of H2 and CO2 to CH4 and acetate in fed-batch biogas reactors by mixed biogas community: a novel route for the power-to-gas concept. Biotechnol. Biofuels 9, 102. https://doi.org/10.1186/s13068-016-0515-0

Szuhaj, M., Wirth, R., Bagi, Z., Maróti, G., Rákhely, G., Kovács, K.L., 2021. Development of Stable Mixed Microbiota for High Yield Power to Methane Conversion. Energies 14, 7336. https://doi.org/10.3390/en14217336

Tamames, J., Puente-Sánchez, F., 2019. SqueezeMeta, a highly portable, fully automatic metagenomic analysis pipeline. Front. Microbiol. https://doi.org/10.3389/fmicb.2018.03349

Tjørve, K.M.C., Tjørve, E., 2017. The use of Gompertz models in growth analyses, and new Gompertz-model approach: An addition to the Unified-Richards family. PLoS One 12, 1–17. https://doi.org/10.1371/journal.pone.0178691

Treu, L, Campanaro, S., Kougias, P.G., Sartori, C., Bassani, I., Angelidaki, I., 2018. Hydrogen-fueled microbial pathways in biogas upgrading systems revealed by genome-centric metagenomics. Front. Microbiol. 9, 1079. https://doi.org/10.3389/fmicb.2018.01079

Treu, L., Campanaro, S., Kougias, P.G., Zhu, X., Angelidaki, I., 2016a. Untangling the Effect of Fatty Acid Addition at Species Level Revealed Different Transcriptional Responses of the Biogas Microbial Community Members. Environ. Sci. Technol. 50,

6079–6090. https://doi.org/10.1021/acs.est.6b00296

Treu, L., Kougias, P.G., Campanaro, S., Bassani, I., Angelidaki, I., 2016b. Deeper insight into the structure of the anaerobic digestion microbial community; The biogas microbiome database is expanded with 157 new genomes. Bioresour. Technol. 216, 260–266. https://doi.org/10.1016/j.biortech.2016.05.081

Treu, L., Kougias, P.G., de Diego-Díaz, B., Campanaro, S., Bassani, I., Fernández-Rodríguez, J., Angelidaki, I., 2018. Two-year microbial adaptation during hydrogen-mediated biogas upgrading process in a serial reactor configuration. Bioresour. Technol. 264, 140–147. https://doi.org/10.1016/j.biortech.2018.05.070

Tsapekos, P., Kougias, P.G., Angelidaki, I., 2018. Mechanical pretreatment for increased biogas production from lignocellulosic biomass; predicting the methane yield from structural plant components. Waste Manag. https://doi.org/10.1016/j.wasman.2018.07.017

Tubeileh, A., Rennie, T.J., Goss, M.J., 2016. A review on biomass production from C4 grasses: yield and quality for end-use. Curr. Opin. Plant Biol. 31, 172–180. https://doi.org/10.1016/J.PBI.2016.05.001

Uellendahl, H., Wang, G., Møller, H.B., Jørgensen, U., Skiadas, I. V., Gavala, H.N., Ahring, B.K., 2008. Energy balance and cost-benefit analysis of biogas production from perennial energy crops pretreated by wet oxidation. Water Sci. Technol. 58, 1841–1847. https://doi.org/10.2166/wst.2008.504

Uritskiy, G. V., DiRuggiero, J., Taylor, J., 2018. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome 6, 158. https://doi.org/10.1186/s40168-018-0541-1

Valentine, J., Clifton-Brown, J., Hastings, A., Robson, P., Allison, G., Smith, P., 2012. Food vs. fuel: The use of land for lignocellulosic “next generation” energy crops that minimize competition with primary food production. GCB Bioenergy. https://doi.org/10.1111/j.1757-1707.2011.01111.x

van der Lelie, D., Taghavi, S., McCorkle, S.M., Li, L.-L., Malfatti, S.A., Monteleone, D.,

Donohoe, B.S., Ding, S.-Y., Adney, W.S., Himmel, M.E., Tringe, S.G., 2012. The metagenome of an anaerobic microbial community decomposing poplar wood chips. PLoS One 7, e36740. https://doi.org/10.1371/journal.pone.0036740

Vanwonterghem, I., Jensen, P.D., Rabaey, K., Tyson, G.W., 2015. Temperature and solids retention time control microbial population dynamics and volatile fatty acid production in replicated anaerobic digesters. Sci. Reports 2015 51 5, 1–8. https://doi.org/10.1038/srep08496

Volk, T.A., Heavey, J.P., Eisenbies, M.H., 2016. Advances in shrub-willow crops for bioenergy, renewable products, and environmental benefits. Food Energy Secur. 5, 97–106. https://doi.org/10.1002/fes3.82

Wagner, A.O., Lackner, N., Mutschlechner, M., Prem, E.M., Markt, R., Illmer, P., 2018. Biological pretreatment strategies for second-generation lignocellulosic resources to enhance biogas production. Energies. https://doi.org/10.3390/en11071797

Wahid, R., Mulat, D.G., Gaby, J.C., Horn, S.J., 2019. Effects of H2:CO2 ratio and H2 supply fluctuation on methane content and microbial community composition during in-situ biological biogas upgrading. Biotechnol. Biofuels 12, 1–15. https://doi.org/10.1186/s13068-019-1443-6

Wang, J.P., Matthews, M.L., Williams, C.M., Shi, R., Yang, C., Tunlaya-Anukit, S., Chen, H.-C., Li, Q., Liu, J., Lin, C.-Y., Naik, P., Sun, Y.-H., Loziuk, P.L., Yeh, T.-F., Kim, H., Gjersing, E., Shollenberger, T., Shuford, C.M., Song, J., Miller, Z., Huang, Y.-Y., Edmunds, C.W., Liu, B., Sun, Y., Lin, Y.-C.J., Li, W., Chen, H., Peszlen, I., Ducoste, J.J., Ralph, J., Chang, H.-M., Muddiman, D.C., Davis, M.F., Smith, C., Isik, F., Sederoff, R., Chiang, V.L., 2018. Improving wood properties for wood utilization through multi-omics integration in lignin biosynthesis. Nat. Commun. 2018 91 9, 1–16. https://doi.org/10.1038/s41467-018-03863-z

Wannasek, L., Ortner, M., Kaul, H.-P., Amon, B., Amon, T., 2019. Double-cropping systems based on rye, maize and sorghum: Impact of variety and harvesting time on biomass and biogas yield. Eur. J. Agron. 110, 125934. https://doi.org/10.1016/j.eja.2019.125934

Waters, C.N., Zalasiewicz, J., Summerhayes, C., Barnosky, A.D., Poirier, C., Gałuszka, A., Cearreta, A., Edgeworth, M., Ellis, E.C., Ellis, M., Jeandel, C., Leinfelder, R., McNeill, J.R., Richter, D. deB., Steffen, W., Syvitski, J., Vidas, D., Wagreich, M., Williams, M., Zhisheng, A., Grinevald, J., Odada, E., Oreskes, N., Wolfe, A.P., 2016. The Anthropocene is functionally and stratigraphically distinct from the Holocene. Science (80-.). 351.

Weiland, P., 2010. Biogas production: Current state and perspectives. Appl. Microbiol. Biotechnol. https://doi.org/10.1007/s00253-009-2246-7

Westerholm, M., Castillo, M. d. P., Chan Andersson, A., Jahre Nilsen, P., Schnürer, A., 2019. Effects of thermal hydrolytic pre-treatment on biogas process efficiency and microbial community structure in industrial- and laboratory-scale digesters. Waste Manag. 95, 150–160. https://doi.org/10.1016/j.wasman.2019.06.004

Whittaker, C., Yates, N.E., Powers, S.J., Misselbrook, T., Shield, I., 2018. Dry matter losses and quality changes during short rotation coppice willow storage in chip or rod form. Biomass and Bioenergy 112, 29–36. https://doi.org/10.1016/J.BIOMBIOE.2018.02.005

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the Tidyverse. J. Open Source Softw. 4, 1686. https://doi.org/10.21105/joss.01686

Wirth, Roland, Kovács, E., Maráti, G., Bagi, Z., Rákhely, G., Kovács, K.L., 2012. Characterization of a biogas-producing microbial community by short-read next generation DNA sequencing. Biotechnol. Biofuels 5, 1–16. https://doi.org/10.1186/1754-6834-5-41

Wirth, R., Kovacs, E., Martín, G., Bagi, Z., Rakhely, G., Kovacs, K.L., 2012. Characterization of a biogas producing microbial community by short read next generation DNA sequencing. Biotechnol Biofuels 5, 1–16. https://doi.org/10.1186/1754-6834-5-41

Wirth, R., Lakatos, G., Böjti, T., Maróti, G., Bagi, Z., Kis, M., Kovács, A., Ács, N., Rákhely, G., Kovács, K.L., 2015. Metagenome changes in the mesophilic biogas-producing community during fermentation of the green alga Scenedesmus obliquus. J. Biotechnol. 215, 52–61. https://doi.org/10.1016/j.jbiotec.2015.06.396

Wirth, R., Lakatos, G., Böjti, T., Maróti, G., Bagi, Z., Rákhely, G., Kovács, K.L., 2018. Anaerobic gaseous biofuel production using microalgal biomass – A review. Anaerobe. https://doi.org/10.1016/j.anaerobe.2018.05.008

Woese, C.R., Fox, G.E., 1977. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. Proc. Natl. Acad. Sci. 74, 5088–5090. https://doi.org/10.1073/pnas.74.11.5088

Woraruthai, T., Kunno, J., Pongsopon, M., Yansakon, K., Phoopraintra, P., Chantiwas, R., Leartsakulpanich, U., Chaiyen, P., Wongnate, T., 2020. Identification and cultivation of hydrogenotrophic methanogens from palm oil mill effluent for high methane production. Int. J. Energy Res. 44, 10058–10070. https://doi.org/10.1002/er.5618

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, 605–607. https://doi.org/10.1093/bioinformatics/btv638

Xu, H., Wang, K., Zhang, X., Gong, H., Xia, Y., Holmes, D.E., 2020. Application of in-situ H2-assisted biogas upgrading in high-rate anaerobic wastewater treatment. Bioresour. Technol. 299, 122598. https://doi.org/10.1016/j.biortech.2019.122598

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.H., Whitman, W.B., Euzéby, J., Amann, R., Rosselló-Móra, R., 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat. Rev. Microbiol. 12, 635–645. https://doi.org/10.1038/nrmicro3330

Yenigün, O., Demirel, B., 2013. Ammonia inhibition in anaerobic digestion: A review. Process Biochem. 48, 901–911. https://doi.org/10.1016/j.procbio.2013.04.012

Yu, G., Wang, L.G., Han, Y., He, Q.Y., 2012. ClusterProfiler: An R package for comparing biological themes among gene clusters. Omi. A J. Integr. Biol. 16, 284–287.

https://doi.org/10.1089/omi.2011.0118

Zhang, L., Kuroki, A., Tong, Y.W., 2020. A Mini-Review on In situ Biogas Upgrading Technologies via Enhanced Hydrogenotrophic Methanogenesis to Improve the Quality of Biogas From Anaerobic Digesters. Front. Energy Res. 8, 1–7. https://doi.org/10.3389/fenrg.2020.00069

Zhao, X., Liu, Jinhuan, Liu, Jingjing, Yang, F., Zhu, W., Yuan, X., Hu, Y., Cui, Z., Wang, X., 2017. Effect of ensiling and silage additives on biogas production and microbial community dynamics during anaerobic digestion of switchgrass. Bioresour. Technol. 241, 349–359. https://doi.org/10.1016/j.biortech.2017.03.183

Zhu, X., Campanaro, S., Treu, L., Kougias, P.G., Angelidaki, I., 2019. Novel ecological insights and functional roles during anaerobic digestion of saccharides unveiled by genome-centric metagenomics. Water Res. 151, 271–279. https://doi.org/10.1016/j.watres.2018.12.041

Zhu, X., Campanaro, S., Treu, L., Seshadri, R., Ivanova, N., Kougias, P.G., Kyrpides, N., Angelidaki, I., 2020. Metabolic dependencies govern microbial syntrophies during methanogenesis in an anaerobic digestion ecosystem. Microbiome 8, 22. https://doi.org/10.1186/s40168-019-0780-9

10. Summary

The harmful impact of humanity on the Earth system includes the increasing rate of species extinction, deforestation, climate change, environmental pollution and overpopulation. One central issue is the increasing global energy demand, which is still met with 80% by fossil fuels. The burning of these contributes to many of the beforementioned global issues, as they produce green-house-gases (GHG) and pollutants. Fortunately, reducing their share is the target of several international treaties. Renewable energies provide an alternative to this, but their combined usage is required to provide effective GHG-mitigation. Biomass is a ubiquitous renewable energy source that can be harnessed in many ways. Because its many advantages, one very promising technology for alternative energy generation from biomass is the anaerobic digestion process (AD). In the context of climate change, the generation of biogas as a form of renewable energy has become popular and was intensively examined over the last few decades. In the course of my doctoral work, I aimed to contribute to the following two major challenges that biogas industry faces:

1.) To find novel substrates that have high biogas production potential, yet cheap to produce and do not come into contradiction with the food vs feed debate, as traditional substrates, such as maize do. Willow biomass may be a promising choice, since producing it has many advantages, including low production costs and the possibility of growing it on marginal lands, but the high amounts of lignin in the woody biomass hindered its usage as a substrate for biogas generation until now. To this end, I utilized a novel lignocellulosic biogas substrate: the young shrubs of the short-rotation coppice willow (Salix spp.), which I termed Green Willow Biomass (GWB).
2.) The coupling of other renewable sources, such as photovoltaics and wind energy, with biogas to harness their excess energy produced when the energy consumption rate is lower than its generation. This can be achieved with a process called a power-to-methane (P2bioCH_4), wherein the excess energy is used to generate H_2 via water electrolysis, which in turn is fed to an anaerobic digester to increase the methane

concentration of the produced biogas. To better understand this *in situ* biological biogas upgrading process (BBU), I investigated the near-immediate effects of hydrogen-addition on the biogas producing microbial community via a metatranscriptomic approach.

Within the first part of my research, I analyzed the fermentation efficiency of GWB, based on parameters, such as maximal methane yields (K) and highest methane production rates (*μmax*). This showed that both the commercial willow genotype (EN) and the novel tetraploid variant (PP) can be viable alternative AD feedstocks. Perhaps even more so when considering CSTR fermentations and industrial-scales operations because the GWB's higher *μmax* values, compared to woody willow (WWB). In addition to the laboratory experiments, I estimated the biomass yields (t/ha) based on a small-scale field experiment to compare the energy potential of GWB with other feedstocks. When considering biomass yields, GWB can be an even more promising choice, as the extractable energy potential (GJ/ha) in the August harvests of GWB almost reached and in an estimated September harvest, it can potentially surpass the corresponding yields of either maize silage or combusted woody willow. Many of these advantages come from the presence (or rather, high amounts) of leaf biomass. Therefore, my thesis provides substantial support to promote and recommend green willow biomass as valuable feedstock for biomethanation.

Within the course of the second part of my thesis, I examined the early response of the biogas-producing microbial community to the presence of saturating amount of H_2. I used a combined metagenomic and metatranscriptomic analysis to determine the changes of the expression levels of the various genes of the metagenome-assembled genomes (MAGs), with a focus to those related to methanogenesis. My results showed that the microbial community (both the *Archaea* and many *Bacteria*) responded practically instantaneously (2 hours after the induced H_2) to the changed environment. The activity of pluripotential methanogens, mainly *Methanosarcina* MAGs, reduced significantly, while the hydrogenotrophic *Methanobacter* MAGs increased. Many MAGs were unique to my samples, but these were not always just side-players, rather several of them are abundant,

main-players of the community. Thus, it seems that specific microbes proliferate in the samples, which is likely the case in ultimately any AD environment. This thorough analysis of the differences between the H_2-treated metatranscriptomes and corresponding controls identified the early events in the microbial communities brought about by the H_2 addition, contributing to our understanding of the BBU mechanism in a P2bioCH4 system.

11. Summary in Hungarian (Összefoglaló)

Az emberiség Földünknek és a bioszférának egyre nagyobb és egyre inkább visszafordíthatatlan károkat okozott. Ennek egyik oka a folyamatosan növekedő energiaigényünk, hiszen kielégítésére leginkább fosszilis energiahordozókat égetünk el, azonban ezáltal üvegházhatású gázok és különböző károsanyagok szabadulnak fel, jutnak környezetünkbe. A megújuló energiák használata részben megoldást nyújthat, ám szükség van ezen technológiák továbbfejlesztésére, hogy versenyezni tudjanak a fosszilis energiákkal. A biomassza az egyik legnagyobb mértékben hozzáférhető energiaforrás, melyből számos energiahordozó állítható elő. Az egyik legszélesebb körben felhasználható ezek közül a biogáz, amely a biomassza levegőmentes bomlása (anaerob fermentáció) során keletkező, metánt és szén-dioxidot tartalmazó gázelegy. A biogáz előállítása rendkívül sokféle alapanyagból, vagyis szubsztrátból lehetséges, mint pl. szennyvíziszap, mezőgazdasági és élelmiszeripari melléktermékek, sertés- vagy marhatrágya, illetve a leggyakrabban használt energianövények. A technológia másik előnye, hogy maga a termék is, tehát a biogáz, különböző módokon felhasználható: elégethető gázmotorokban elektromosság előállítására, felhasználható otthonok fűtésére vagy a közlekedésben. Legnagyobb potenciáljának kihasználására akkor nyílik alkalom, ha adekvát tisztítás által biometánt állítunk elő belőle, amely a földgázzal ekvivalens, így lehetséges a földgázhálózatba betáplálni, bár ez Magyarországon még csak 1 esetben valósult meg.

Doktori disszertációm során a biogáztermelés két központi problémájának megoldásához szerettem volna hozzájárulni:

1.) Jelenleg a biogáztermelés leginkább tradicionális energianövények (kukorica, szója) felhasználásával valósul meg, ám egyre inkább nő az igény olyan potenciális alapanyagok iránt, amelyek termesztését meg lehet valósítani marginális területeken, ezáltal nem kerülnek összetűzésbe élelmezési célokra termesztett növényekkel (pl. búza, kukorica). Ilyen lignocellulóz-alapú szubsztrát lehet a rövid-vetésfordulójú energiafűz (RVF-EF) is. Az RVF-EF-et azonban eddig csak biomassza-erőművekben történő égetés céljából termesztették; a belőle történő

biogáz-előállításra ugyan volt már példa a szakirodalomban, de csak a magas lignocellulóz-tartalom miatt szükséges előkezelések elvégzése után volt alkalmas erre a célra. Az általam javasolt rendszerben az eddigiekkel ellentétben az RVF-EF fiatalon, még zöld állapotában kerül begyűjtésre, hogy a nagymennyiségű levél-biomassza is felhasználásra kerüljön. Az így kapott biomasszát Zöld Fűz Biomasszának (ZFB-nek) neveztem el. A disszertációm első fő részében megvizsgáltam ezen rendszer használatának lehetőségét.

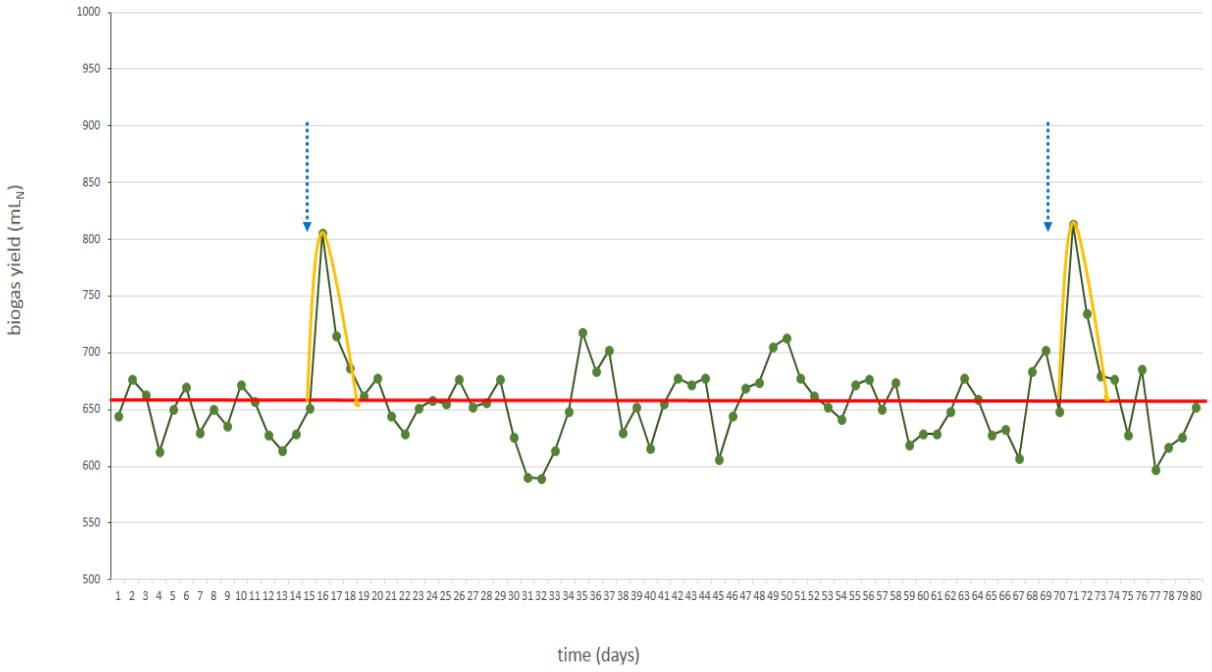
2.) A biometán előállításának potenciális lehetőségét egy *in situ* biológiai biogázjavítással kapcsolt power-to-gas (P2G, „áram → gáz") rendszerben, ahol az egyéb megújuló forrásokból termelt fel nem használt elektromos energiából első lépésben H_2 előállítása történik meg (víz elektrolízis segítségével), majd az ez kerül betáplálásra az utófermentációs tartályba, vagy akár közvetlenül fermentációs egységekbe. Ez az ún. power-to-biomethane (P2bioCH_4, „áram → biometán") koncepció, amelyben a H_2 központi szerepet játszik, azonban annak a biogáztermelő mikroközösségre gyakorolt rövidtávú hatását még nem vizsgálták. Dolgozatom második felében ezért ezt a hatást vizsgáltam meg genom-centrikus metatranszkriptomikai módszer segítségével.

A ZFB vizsgálatának során egy kisméretű, 100 növényből álló kísérleti ültetvény biomasszáját használtam fel. A kísérleti populáció 2 genotípusú növényből állt, a betakarítást 4 alkalommal végeztem el. Megmértem a szárak és a levelek mennyiségét, biometán hozamukra illesztett növekedési modell paramétereit (K=teljes metánhozam, μmax=maximális metántermelési ráta) szakaszos üzemű anaerob reaktorokban. Az összehasonlításból megállapítottam, hogy a legoptimálisabb modell paramétereket a júniusi betakarítás mutatta, valamint, hogy mindkét genotípus alkalmas lehet biogáz fermentorokban történő felhasználásra, valamint, hogy a levélrész jelentősen magasabb μmax értékeknek köszönhetően gyorsabban degradálódik a szárnál, ami különösen előnyös folyamatos üzemű fermentációk esetében. Amennyiben számításba vesszük a biomasszahozamokat is megállapítható, hogy egy késő augusztusi-kora szeptemberi

betakarítás lenne a legelőnyösebb, hiszen az ültetvény ilyenkor mutatja a legmagasabb energiapotenciált (amelyet a metánhozam és a biomasszahozam összeszorzásából kaphatunk meg). Ezek az értékek megközelítik, sőt potenciálisan meg is haladhatják a silókukorica értékeit (amennyiben azt szintén biogáz-előállításra használnánk), valamint a fás fűz értékeit, amennyiben az a tradicionális módon biomassza-égetőkben kerülne felhasználásra. Mindezek arra mutatnak, hogy az energiafűz egyéb előnyei mellett, melyeket más energianövényekkel szemben mutat, a ZFB felhasználása biogázüzemekben egy igen gazdaságos megoldás lehet.

Dolgozatom második felében megvizsgáltam az anaerob reaktorokba történő H_2-betáplálás korai hatását a mikrobaközösségre, központi figyelmet szentelve a metántermelő Archaeák (metanogének) populációinak génkifejeződésének változásaira. A kísérlet során egy stabil biogáztermelő folyamatos üzemű anaerob reaktor szolgált modellként az „áram → biometán" koncepció vizsgálatához. Megállapítottam, hogy már 2 óra elteltével is mind az Archaeális, mind a Bakteriális populációk génexpressziós profilja, illetve metabolikus útvonalainak kifejeződésének mértéke jelentősen változott a H_2 hatására. A pluripotens metanogének, leginkább *Methanosarcinák* abundanciája jelentősen csökkent, míg a hidrogenotróf metanogének (*Methanobacter*) aktivizálódtak. A kísérleti minták biológiai párhuzamosai 2 hónap elteltével is nagyfokú hasonlóságot mutattak, ami bizonyítéka nem csak a felhasznált módszerek robusztusságának, hanem magának a rendszer plaszticitásának a H2 adagolás ki/be-kapcsolásával szemben is. Ezen *in situ* biológiai biogázjavítási eljárás tehát egy ígéretes módszer lehet az „áram → biometán" rendszer által történő bioenergia felhasználásra.

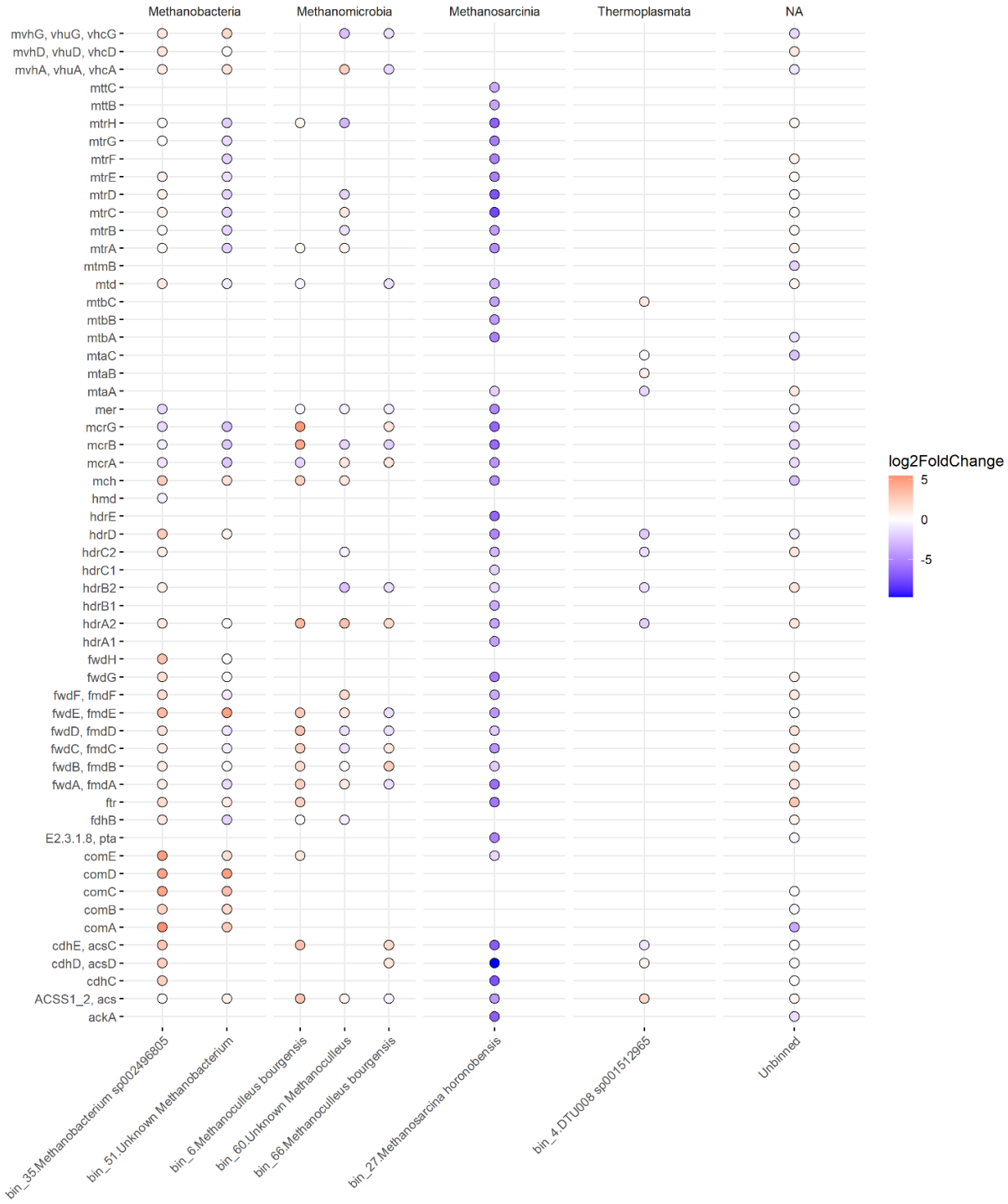
12. Supplementary Material



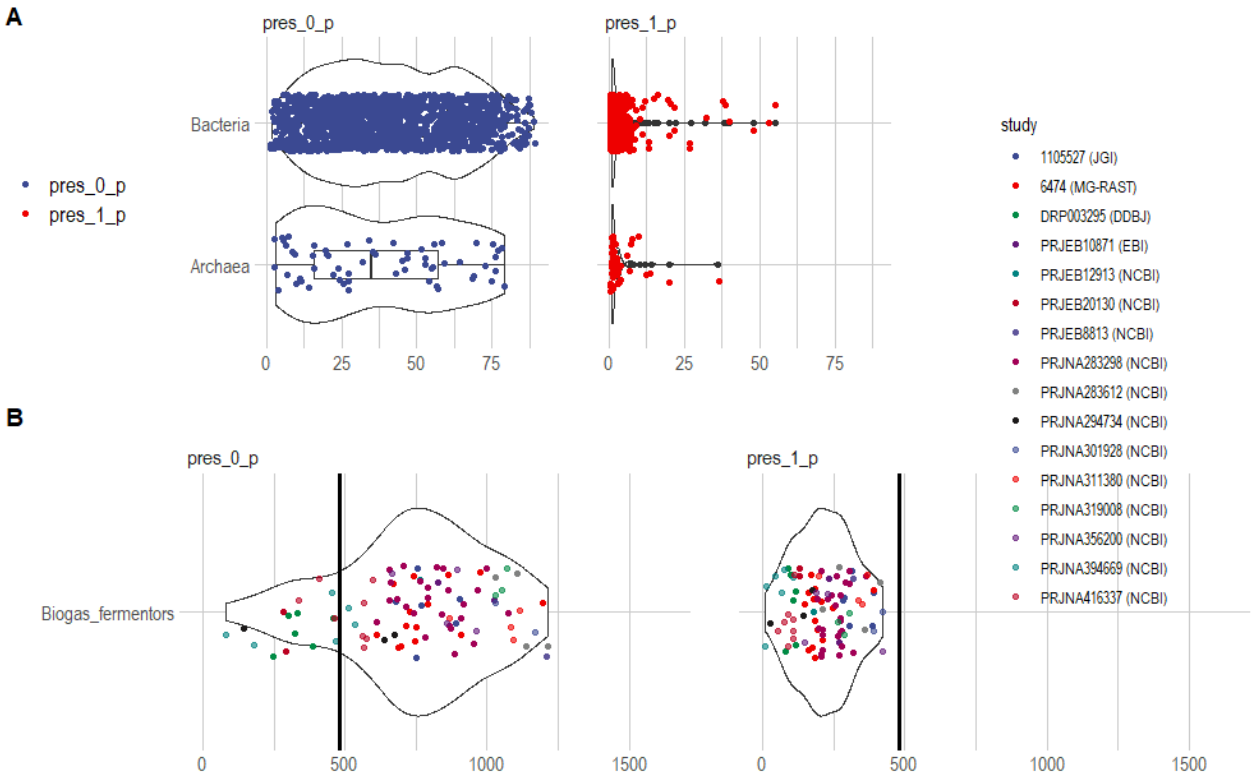
Supplementary Figure S1. Biogas yields during the course of the H_2-addition experiment. Blue arrows indicate H_2 injections.

Bin statistics			Taxonomy based on GTDB database (https://gtdb.ecogenomic.org/)							biome (http://microbial-g	
Bin nar	Com	Redu	domain	phylum	class	order	family	genus	species	Closest relat	Percent A
Bin_1	97.18	7.04	Bacteria	Firmicutes	Clostridia	Acetivibrionales	Acetivibrionaceae	Herbivorax	Herbivorax saccincola	#	0
Bin_4	64.47	0	Archaea	Thermoplasmatot	Thermoplasmata	Methanomassiliicocc	Methanomassiliicocc	DTU008	DTU008 sp001512965	#	0
Bin_6	72.37	11.84	Archaea	Halobacterota	Methanomicrobia	Methanomicrobiales	Methanocullaceae	Methanoculleus	Methanoculleus bourgensis	#	0
Bin_18	60.53	0	#	#	#	#	#	#	#	#	0
Bin_20	69.01	12.68	Bacteria	Firmicutes	Mahellia	Caldicoprobacterales	Caldicoprobacteracea	DTU074	DTU074 sp002385885	#	0
Bin_21	76.06	16.9	Bacteria	Firmicutes	Thermovenabulia	Thermovenabulales	Tepidanaerobacterac	DTU063	DTU063 sp001512695	#	0
Bin_34	33.8	2.82	Bacteria	Hydrogenedentot	Hydrogenedentia	Hydrogenedentiales	UBA2224	UBA2224	UBA2224 sp002348185	#	0
Bin_35	97.37	19.74	Archaea	Euryarchaeota	Methanobacteria	Methanobacteriales	Methanobacteriaceae	Methanobacterium	Methanobacterium sp002496	#	0
Bin_52	35.21	0	Bacteria	Bacteroidota	Bacteroidia	Bacteroidales	Marinilabiliaceae	UBA7646	UBA7646 sp002411385	#	0
Bin_54	42.25	1.41	Bacteria	Fibrobacterota	Chitinivibrionia	Chitinivibrionales	Chitinispirillaceae	Chitinispirillum	Chitinispirillum alkaliphilum	#	0
Bin_58	91.55	54.93	Bacteria	Firmicutes	#	#	#	#	#	#	0
Bin_59	71.83	21.13	Bacteria	Firmicutes	Limnochordia	DTU010	DTU010	DTU010	DTU010 sp900018335	#	0
Bin_60	75	6.58	Archaea	Halobacterota	Methanomicrobia	Methanomicrobiales	Methanocullaceae	Methanoculleus		#	0
Bin_66	82.89	3.95	Archaea	Halobacterota	Methanomicrobia	Methanomicrobiales	Methanocullaceae	Methanoculleus	Methanoculleus bourgensis	#	0
Bin_31	66.2	2.82	#	#	#	#	#	#	#	METABAT AS06	34.48
Bin_48	47.89	0	#	#	#	#	#	#	#	METABAT AS27	43.95
Bin_39	46.48	1.41	Bacteria	Firmicutes	DTU014	DTU014	DTU014	DTU014	DTU014 sp900016865	METABAT AS04	46.44
Bin_37	87.32	0	#	#	#	#	#	#	#	METABAT AS22	50.22
Bin_24	57.75	0	Bacteria	Firmicutes	Bacilli	Haloplasmatales	Turicibacteraceae	Turicibacter	Turicibacter sp001543345	METABAT AS04	51.86
Bin_55	46.48	0	Bacteria	Firmicutes	Clostridia	Lachnospirales	Cellulosilyticaceae	Cellulosilyticum	#	METABAT AS21	54.83
Bin_47	26.76	0	#	#	#	#	#	#	#	METABAT AS05	56.39
Bin_51	76.32	0	Archaea	Euryarchaeota	Methanobacteria	Methanobacteriales	Methanobacteriaceae	Methanobacterium	#	METABAT AS08	56.88
Bin_19	88.73	1.41	Bacteria	Chloroflexota	Anaerolineae	Anaerolineales	Anaerolineaceae	T78	T78 sp002305635	METABAT AS06	63.25
Bin_27	92.11	0	Archaea	Halobacterota	Methanosarcinia	Methanosarcinales	Methanosarcinaceae	Methanosarcina	Methanosarcina horonobensi	METABAT AS10	70.69
Bin_49	42.25	0	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	Clostridium	Clostridium sp001916075	METABAT AS07	73.41
Bin_45	32.39	0	#	#	#	#	#	#	#	METABAT AS07	76.51
Bin_71	46.48	0	Bacteria	Bacteroidota	Bacteroidia	Bacteroidales	UBA932	Bact-11	#	METABAT AS06	76.92
Bin_73	71.83	0	#	#	#	#	#	#	#	METABAT AS10	78.3
Bin_57	77.46	0	#	#	#	#	#	#	#	METABAT AS21	95.41
Bin_16	67.61	1.41	Bacteria	Firmicutes	Clostridia	Oscillospirales	Oscillospiraceae	UBA5446	UBA5446 sp002427465	METABAT AS4L	96.17
Bin_30	64.79	2.82	Bacteria	Patescibacteria	Paceibacteria	Paceibacterales	UBA5633	UBA2558	UBA2558 sp002340425	METABAT AS24	96.87
Bin_42	92.96	0	Bacteria	Synergistota	Synergistia	Synergistales	54-24	54-24	54-24 sp001508935	METABAT AS04	97.09
Bin_70	73.24	1.41	Bacteria	Actinobacteriota	Actinobacteria	Mycobacteriales	Mycobacteriaceae	Corynebacterium	Corynebacterium humireduce	METABAT AS15	97.19
Bin_79	87.32	0	Bacteria	Cloacimonadota	Cloacimonadia	Cloacimonadales	Cloacimonadaceae	UBA3900	UBA3900 sp002391675	METABAT AS26	97.42
Bin_26	87.32	1.41	Bacteria	Cloacimonadota	Cloacimonadia	Cloacimonadales	Cloacimonadaceae	UBA4175	UBA4175 sp002379855	METABAT AS4B	97.59
Bin_17	77.46	0	#	#	#	#	#	#	#	METABAT AS16	97.76
Bin_83	100	0	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	#	METABAT AS08	98.1
Bin_65	57.75	0	#	#	#	#	#	#	#	METABAT AS22	98.14
Bin_5	88.73	2.82	Bacteria	Bacteroidota	Bacteroidia	Bacteroidales	Bacteroidaceae	UBA1179	UBA1179 sp002340405	METABAT AS10	98.19
Bin_56	50.7	1.41	Bacteria	Firmicutes	Clostridia	Lachnospirales	Defluviitaleaceae	Defluviitalea	Defluviitalea phaphyphila	METABAT AS23	98.27
Bin_40	81.69	1.41	Bacteria	Spirochaetota	Spirochaetia	Treponematales	Treponemataceae	Treponema	Treponema brennaborense	METABAT AS23	98.38
Bin_23	91.55	7.04	Bacteria	Firmicutes	Bacilli	#	#	#	#	METABAT AS23	98.39
Bin_33	57.75	0	Bacteria	Firmicutes	Bacilli	RF39	UBA5348	UBA5348	UBA5348 sp002411005	METABAT AS08	98.71
Bin_44	90.14	1.41	Bacteria	Firmicutes	Clostridia	Saccharofermentanal	DTU023	UBA4923	UBA4923 sp002427535	METABAT AS21	98.76
Bin_38	92.96	0	#	#	#	#	#	#	#	METABAT AS15	98.86
Bin_28	91.55	2.82	Bacteria	Spirochaetota	Spirochaetia	Treponematales	Treponemataceae	Spiro-10	Spiro-10 sp001604405	METABAT AS22	98.88
Bin_80	59.15	1.41	Bacteria	Firmicutes	Clostridia	Lachnospirales	Lachnospiraceae	Herbinix	Herbinix luporum	METABAT AS23	98.89
Bin_46	50.7	2.82	Bacteria	Firmicutes	Clostridia	Acetivibrionales	Ruminiclostridiaceae	Ruminiclostridium	Ruminiclostridium sp001512	METABAT AS06	98.9
Bin_62	84.51	2.82	#	#	#	#	#	#	#	METABAT AS22	98.96
Bin_61	56.34	4.23	Bacteria	Firmicutes	Limnochordia	DTU010	DTU010	DTU010	DTU010 sp002391385	METABAT AS24	99.01
Bin_36	49.3	0	Bacteria	Bacteroidota	Bacteroidia	Bacteroidales	vadinHA17	SR-FBR-E99	SR-FBR-E99 sp002409145	METABAT AS05	99.06
Bin_12	80.28	0	Bacteria	Firmicutes	Bacilli	RFN20	CAG-826	UBA1361	UBA1361 sp002306335	METABAT AS24	99.11
Bin_32	46.48	0	Bacteria	Bacteroidota	Bacteroidia	Bacteroidales	WCHB1-69	UBA5429	UBA5429 sp002427605	METABAT AS08	99.11
Bin_75	78.87	5.63	Bacteria	Firmicutes	Clostridia	Peptostreptococcales	Anaerovoracaceae	UBA1426	UBA1426 sp002385825	METABAT AS21	99.11
Bin_25	74.65	2.82	#	#	#	#	#	#	#	METABAT AS08	99.13
Bin_63	84.51	0	Bacteria	Firmicutes	Mahellia	Caldicoprobacterales	DTU083	UBA3906	UBA3906 sp002391555	METABAT AS2H	99.13
Bin_22	94.37	0	#	#	#	#	#	#	#	METABAT AS08	99.14
Bin_9	77.46	2.82	Bacteria	Bacteroidota	Bacteroidia	Bacteroidales	ML635J-15	Bact-19	Bact-19 sp002412425	METABAT AS24	99.18
Bin_76	94.37	1.41	Bacteria	Caldatribacteriota	Caldatribacteriia	Caldatribacteriales	Caldatribacteriaceae	UBA3950	UBA3950 sp002385475	METABAT AS07	99.27
Bin_43	95.77	1.41	Bacteria	Verrucomicrobiot	Kiritimatiellae	RFP12	UBA1067	UBA1211	UBA1211 sp002309585	METABAT AS08	99.31
Bin_78	78.87	0	#	#	#	#	#	#	#	METABAT AS4G	99.41
Bin_82	85.92	1.41	Bacteria	Firmicutes	SHA-98	UBA4971	UBA4971	UBA2557	UBA2557 sp900019985	METABAT AS07	99.44
Bin_50	87.32	0	#	#	#	#	#	#	#	METABAT AS24	99.48
Bin_2	43.66	1.41	Bacteria	Synergistota	Synergistia	Synergistales	Synergistaceae	Syner-03	Syner-03 sp002306075	METABAT AS21	99.49
Bin_69	84.51	0	Bacteria	Synergistota	Synergistia	Synergistales	Acetomicrobiaceae	Acetomicrobium	Acetomicrobium flavidum	METABAT AS24	99.49
Bin_68	91.55	1.41	Bacteria	Firmicutes	Clostridia	Acetivibrionales	Acetivibrionaceae	Ruminiclostridium	Ruminiclostridium sp001512	METABAT AS24	99.51
Bin_72	52.11	1.41	Bacteria	Firmicutes	DTU065	DTU065	DTU065	DTU065	DTU065 sp001512545	METABAT AS02	99.52
Bin_15	94.37	0	Bacteria	Firmicutes	Clostridia	Acetivibrionales	Acetivibrionaceae	DTU013	DTU013 sp002385815	METABAT AS24	99.54
Bin_8	77.46	0	Bacteria	Bacteroidota	Bacteroidia	Bacteroidales	Dysgonomonadaceae	Fermentimonas	Fermentimonas massiliensis	METABAT AS08	99.56
Bin_11	84.51	2.82	Bacteria	Firmicutes	Mahellia	Caldicoprobacterales	UBA3941	UBA3941	UBA3941 sp002385665	METABAT AS25	99.56
Bin_67	97.18	0	Bacteria	Firmicutes	Clostridia	Tissierellales	Sedimentibacteracea	Sedimentibacter	Sedimentibacter sp002316225	METABAT AS21	99.62
Bin_64	91.55	0	Bacteria	Verrucomicrobiot	Kiritimatiellae	LD1-PB3	Lenti-01	Lenti-01	Lenti-01 sp002304915	METABAT AS08	99.63
Bin_77	94.37	0	Bacteria	Bacteroidota	Bacteroidia	Bacteroidales	DTU049	DTU049	DTU049 sp001512885	METABAT AS26	99.67
Bin_14	40.85	2.82	Bacteria	Bacteroidota	Bacteroidia	Bacteroidales	WCHB1-69	UBA1205	UBA1205 sp002309615	METABAT AS22	99.68
Bin_3	77.46	1.41	Bacteria	Bacteroidota	Bacteroidia	Bacteroidales	WCHB1-69	UBA5429	UBA5429 sp002427605	METABAT AS08	99.69
Bin_84	87.32	0	Bacteria	Firmicutes	Clostridia	Acetivibrionales	Ruminiclostridiaceae	DTU059	DTU059 sp001512555	METABAT AS24	99.69
Bin_7	90.14	1.41	Bacteria	Bacteroidota	Bacteroidia	Bacteroidales	Dysgonomonadaceae	#	#	METABAT AS05	99.71
Bin_74	83.1	0	Bacteria	Firmicutes	Dethiobacteria	Dethiobacterales	Dethiobacteraceae	DTU027	DTU027 sp002385745	METABAT AS08	99.72
Bin_41	64.79	0	Bacteria	Bacteroidota	Bacteroidia	Bacteroidales	Muribaculaceae	CAG-1031	CAG-1031 sp000431215	METABAT AS23	99.74
Bin_53	94.37	7.04	Bacteria	Firmicutes	Desulfotomaculia	Desulfotomaculales	Pelotomaculaceae	DTU098	DTU098 sp001512635	METABAT AS08	99.74
Bin_10	91.55	0	Bacteria	Firmicutes	Clostridia	Lachnospirales	Lachnospiraceae	Herbinix	Herbinix luporum	METABAT AS08	99.79
Bin_81	85.92	0	Bacteria	Firmicutes	Clostridia	Lachnospirales	Lachnospiraceae	Herbinix	Herbinix luporum	METABAT AS24	99.8
Bin_13	98.59	0	Bacteria	Firmicutes	Clostridia	4C28d-15	DTU072	DTU072	DTU072 sp001512685	METABAT AS08	99.9
Bin_29	85.92	0	Bacteria	Firmicutes	Clostridia	4C28d-15	DTU072	DTU072	DTU072 sp001512685	METABAT AS08	99.91

Supplementary Table S2. (previous page). Taxonomic assignments of the MAGs and their comparison to the BGMB



Supplementary Figure S2. Mean *l2FC* of all methanogenesis-related genes of Archaeal MAGs.



Supplementary Figure S3.) A. Boxplot of occurrence of the Bacterial and Archaeal MAGs from the biogas microbiome collection in the 91 metagenome samples analyzed. B. The number of MAGs present in each of the samples. Occurrence/presence is defined as having at least 0.00001% (pres_0_p) or 0.1% (pres_1_p) abundance in both A. and B. The mean occurrence of MAGs above 0.00001% is 757, while those above 0.1% is 209 (thick lines). Data taken from (Campanaro et al., 2020).