Metagenomics survey unravels diversities of biogas’ microbiomes with potential to enhance its’ productivity in Kenya
Muturi S.M.1,2*, Muthui L.W.3, Njogu P.M.4, Onguso J.M.2, Wachira F.N.5, Opiyo S.O.6,7, and Pelle R.3

1Department of Biological Sciences, University of Eldoret, Kenya, 2Institute for Biotechnology Research, Jomo Kenyatta University of Agriculture and Technology, Kenya, 3Biosciences eastern and central Africa - International Livestock Research Institute (BecA-ILRI) Hub, Kenya, 4Institute for Energy and Environmental Technology, Jomo Kenyatta University of Agriculture and Technology, Kenya, 5Department of Life Sciences, South Eastern Kenya University, Kenya, 6OARDC, Molecular and Cellular Imaging Center-Columbus, Ohio state University, USA, and 7The University of Sacread Heart, Gulu, Uganda.

*Corresponding author, email:samuelmuturi37@gmail.com

ABSTRACT
The obstacle to optimal utilization of biogas technology is poor understanding of biogas’ microbiome diversities over a wide geographical coverage. We performed random shotgun sequencing on twelve environmental samples. A randomized complete block design was utilized to assign the twelve biogas reactor treatments to four blocks, within eastern and central regions of Kenya. We obtained 42 million paired-end reads that were annotated against sixteen reference databases using two ENVO ontologies, prior to β-diversities studies. We identified 37 phyla, 65 classes and 132 orders of micro-organisms. Bacteria dominated the microbiome and comprised of 28 phyla, 42 classes and 92 orders, conveying substrate’s versatility in the treatments. Though, fungi and Archaea comprised of only 5 phyla, the fungi were richer; suggesting the importance of hydrolysis and fermentation in biogas production systems. High β-diversity within the taxa was largely linked to communities’ metabolic capabilities. Clostridiales and Bacteroidales, the most prevalent guilds, metabolize organic macromolecules. The identified affiliates of Cytophagales, Alteromonadales, Flavobacteriales, Fusobacteriales, Deferribacterales, Elusimicrobiales, Chlamydiales, Synergistales to mention but few, also catabolize macromolecules into smaller substrates to conserve energy. Furthermore, δ-Proteobacteria, Gloeobacteria and Clostridia affiliates syntrophically regulate P\textsubscript{H2} and reduce metal to provide reducing equivalents. Methanomicrobiales and other Methanomicrobia species were the most prevalence Archaea, converting formate, CO\textsubscript{2(g)}, acetate and methylated substrates into CH\textsubscript{4(g)}. Thermococci, Thermoplasmata and Thermoprotei were among the sulfur and other metal reducing Archaea that contributed to redox balancing and other metabolism within treatments. Eukaryotes, mainly fungi were the least abundant guild, comprised largely Ascomycota and Basidiomycota species. Chytridiomycetes, Blastocladiomycetes and Mortierellomycetes were among the rare species, suggesting their metabolic and substrates limitations. Generally, we observed that environmental and treatment perturbations influenced communities’ abundance, β-diversity and reactor performance largely through stochastic effect. The study of the diversity of the biogas’ microbiomes over wide environmental variables and the productivity of biogas reactor systems has provided insights into better management strategies that may ameliorate biochemical limitations to effective biogas production.

Keywords: Metagenomics, Microbiome, Diversities, Biogas production

Author Summary
The failure of biochemical reactions in biogas producing systems is a common problem and results from poor functioning of the inhabiting micro-organisms. A poor understanding of the global diversities of these micro-organisms and lack of information on the link between
environmental variables, biogas production, and community composition, contrains the
development of strategies that can ameliorate these biochemical issues. We have integrated
sequencing-by-synthesis technology and intensive computational approaches to reveal
metacommunities in the studied reactor treatments. The identified communities were
compared with the treatment’s phenotypic and environmental data in an attempt to fill the
existing knowledge gaps on biogas microbiomes and their production capacities. We present
132 biogas taxonomic profiles systematically and comparatively, linking the abundance with
the identified environmental variables. The local composition of microbiome and variations
in abundance were also linked to the observed differences in biogas productivity, suggesting
the possible cause of the observed variations. The detailed information presented in this study
can aid in the genetic manipulation or formulation of optimal microbial ratios to improve
their effectiveness in biogas production.

Introduction
The optimal utilization of micro-organisms is a critical strategy in the development of new
products and improvement of the existing bioprocesses [1]. One of the widely applied
technologies that utilises micro-organisms is anaerobic digestion (AD) of waste materials.
The process generates biogas and digestate from wastes, the former being an energy carrier
molecule [2]. Over 30,000 industrial installations globally, have been revealed to utilize AD
to convert organic wastes into power, producing up to 10,000 MW [3]. Apart from power
production, stable AD processes provide one of the most efficient environmental models for
organic waste valorization that enables nutrient recovery from the digestate [4], and hence
providing additional economic benefits [5], when compared to other methods of waste
management [6]. Many African countries, including Kenya produce huge amounts of waste
[5] with high biogas generating potential. In spite of the efforts made to exploit this resource
through the Biogas Partnership Programme [7], the biogas systems in African countries
remain fairly rudimentary. The Programmes have installed approximately 60,000 biogas
plants in Africa which include 16,419 in Kenya, 13,584 in Ethiopia, and 13,037; 7518, and
6504 in Tanzania, Burkina Faso and Uganda, respectively [5, 7]. Though Kenya leads in
biogas exploitation in Africa, adoption of the technology remains poor, and is not
commensurate with the country’s energy needs. However, because of the relative popularity
of the technology in the country, compared to its’ counterparts, Kenya was justified as the
appropriate pilot site for biogas studies among other African countries. Similar guidelines
were also followed to select the Kenyan regions in this study.

Overall, it is thought that the poor adoption of the biogas technology in Africa is partly due
to poor knowledge on the microbiomes that mediate the AD process. These processes
comprise thermodynamic unfavorable reactions that contribute to stability and performance
of the treatments. The microbial populations and the reactions involved are highly sensitive to
environmental and reactor operation conditions; traits that prompt the need for development
of sustainable management strategies for effective biogas production. Without such
strategies, a complete breakdown of the whole system occurs. In the event of complete
failure, farmers are forced to abandon the systems or invest in costly restoration processes,
all of which end up reducing the economic and environmental benefits of the AD systems;
thereby constraining efforts towards greening the economy.

Although, previous reports from other continents have indicated dominance of Bacterial and
Archaeal species in tolerating reactor perturbations, limitations of traditional culture-
dependent and independent methods e.g 16S rRNA gene [8-9], still limit the knowledge on
global biogas microbiomes’ diversities within these systems. The 16S rRNA techniques used
thus far provide low resolution [10, 11, 12], and are not robust enough to unravel the high
level of microbial diversity in biogas ecosystems. All these limit the information available on microbial composition and diversity which would be critical for improvement and better management of biogas systems. The 16S rRNA gene does not reveal the full extent of microbial diversity in an ecosystem [13-17] and most current diversity estimates are extrapolations of empirical scaling laws [13], and other theoretical models of biodiversity [14]. The above limitation constrains the capacity to scrutinize and apply the biodiversity scaling laws, macroecological theories and other ecological theories [12-14] in the context of biogas production. This ends up directly limiting the available microbial diversity knowledge which would be useful to develop viable mitigation strategies to address biochemical reaction failures. Despite covering a small fraction of microbial diversities in a dataset, marker gene based studies, annotate the detected species against biological databases that comprise sequences of interest. This approach has the potential to lead to conflicts on the links between microbial populations and reactor performance. Such conflicts were reported by Wittebolle et al. [18] and Konopka et al [19]. Wittebolle et al. considered higher microbial diversity as a reservoir of redundant metabolic pathways that possess desirable traits that respond to reactor pertubations [18] and other environmental conditions. In contrast, Konopka et al. argued that less diverse communities confer system stability by expressing complementary pathways [19] that avoid direct competition over the available resources. These kinds of contrasting views prompted us to ask whether microbial diversity in biogas reactors matters or it is the existence of the core microbiomes that determine the AD stability and maximum biogas production. In an effort to seek answers to these questions, we were provoked to further characterize the biogas system’s micro-organisms using next generation sequencing platforms also termed “massively parallel sequencing” over a wide geographic area. Specifically, the technology employed sequencing-by-synthesis chemistry to generate the dataset. Previously, this technology has enabled massive sequencing of environmental genomes with high quality and accurancy [20]. However, the application of the technology in the biogas systems is still at its infancy. Furthermore, eukaryotes, mainly fungi have normally been neglected by the majority of the studies, yet these species, are also crucial in the metabolism of large macromolecules in the AD systems. Our study was therefore designed to address the limitations enumerated above. We expect our findings to contribute to the formation of testable hypotheses for future application of metacommunity theory on the links between microbial diversity and environmental conditions. Further the study will contribute to a greater understanding of biogas microbiomes for improvement of biogas production.

Results

Annotations against the 16 biological databases

We used Illumina based shotgun metagenomics to identify and compare microbial communities within and among the biogas reactor treatments. In total 12 sludge samples assigned to twelve different treatments operating at a steady-state were assayed. The procedure yielded a total of 44 million raw reads and out of these, 2 million reads were of poor quality. We revealed 24,189,921 reads of 35-151bps from the remaining 42 million reads, which were filtered based on the % G:C content bias and duplicates according to Gomez-Alvarez et al. [21]. Approximately 23,373,513 reads conveying 1,261833 to 3,986675 reads per the treatment were realized (SFig. 1a). Thereafter, we trimmed and assembled the contigs using de novo approach and the obtained high quality scaffolds (3,819,927 scaffolds, 1,119,465,586 bps, with mean average length of 291bps, and mean average standard deviation of 410 bps, STable 1) were annotated against sixteen databases using two ENVO ontologies. Taxonomic assignments conveyed by SEED database were the highest (180, 000 scaffolds), while the assignment of NOG, LSU, RDP and COG databases were the least (<8000 scaffolds). Among the databases, IMG, Genbank and TrEMBL
conveyed equal scaffold annotations (144,000 scaffolds). Other databases annotated the reads as indicated in Fig. 1. Between the two ontologies, the large lake biome revealed the highest annotation compared to the terrestrial biome with a precision of \( P \leq 0.05 \) across the treatments. Furthermore, 39.75% to 46.70% and 53.07% to 60.11% genes were conveyed to known and unknown proteins.

**Annotations against the archived taxonomic profiles**

Taxonomic assignments revealed four broad but relatively simple distributions of microorganisms, with the species of three domains being responsible for methanogenesis. In summary, 2, 617,791 reads hits were generated, with 213681 hits being assigned to Archaea, 2379293 hits to Bacteria and 21886 hits to Eukaryotes (SFig. 1a). Thus, the revealed abundance hits ratio of Bacteria to Archaea to Eukaryotes in the reactor treatments was 109:9:1. The higher abundance of Bacterial species was linked to their metabolic capabilities. Unlike other cellular organisms, Eukaryotes mainly fungi had the least abundances (< 1% of the total hits) among the treatments. Despite these variations of species abundances, we observed consistent microbial richness among the studied treatments. In total 37 phyla, 73 classes and 132 orders were indentified. A majority of the reactor treatments exhibited high community diversity, except for the communities detected in reactor 1, 3 and 6 that were found to cluster partially on the upper right quadrant of the plot (Fig. 2), at a \( P \leq 0.05 \). All the identified phylotypes were further annotated and \( \beta \)-diversity determined in details.

**The identified Bacterial biomes**

We evaluated the variation in abundance and composition of 92 Bacterial domain orders that were assigned to 41 classes and 21 phyla (SFig. 2a, Table 1). Generally we observed significantly higher Bacterial abundances (89-93% of the total hits), compared to other domains. Similarly, significantly high \( \beta \)-diversities of the Bacteria communities were also observed at the phylum level. However, the communities of reactor 2 and 10 (within the same block) and reactor 4 and 9 (on different blocks) were found to reveal low \( \beta \)-diversity (SFig. 2b). In addition, we further observed low diversity variabilities between communities of reactor 3 and 6 and those of reactor 8 and 11 when compared at lower (class and order) taxonomic ranks (SFig. 3a and 3b). Hence, only four out of twelve treatments had high \( \beta \)-diversities, which could indicate genetic plasticity due to species evolution.
Specifically, we revealed δ-Proteobacteria species as the most abundant guild, constituting 9-14% of the total microbial communities. Globally the species were the second most abundant after the Clostridial species that comprised 11-26% of the total communities. The δ-Proteobacteria communities were classified into 7 orders (SFig. 6a). Among them were Syntrophobacteriales that comprised 2-4% of the total identified communities, followed by Desulfuromonadales and Myxococcales species that comprised 2-4% and 0.9-4% of the total microbial compositions respectively. Desulfobacterales and Desulfovibronales communities were found to comprise 1-2% of the total species abundances. Among the least abundant δ-Proteobacterial species were the affiliates of Desulforculales and Bdellovibrionales that comprised only < 1% of the identified total communities. The non-classified reads derived from the class δ-Proteobacteria, also composed < 1% of the total communities. However, the PCoA plots indicated high sample to sample variation of the δ-Proteobacteria communities, although the composition of reactor 4 and 7 were found to be similar (SFig. 6b).
| Phylum                  | S_11.10 | S_9.10 | S_12.10 | S_6.10 | S_8.10 | S_3.10 | S_1 | S_4.10 | S_2.0 | S_5.10 | S_7.10 | S_10.10 |
|------------------------|---------|--------|---------|--------|--------|--------|-----|--------|-------|--------|--------|---------|
| Firmicutes             | 27%     | 34%    | 26%     | 21%    | 25%    | 22%    | 22% | 31%    | 15%   | 30%    | 29%    | 15%     |
| Proteobacteria         | 26%     | 19%    | 20%     | 27%    | 24%    | 27%    | 25% | 19%    | 31%   | 25%    | 21%    | 32%     |
| Bacteroidetes          | 15%     | 18%    | 22%     | 11%    | 18%    | 11%    | 15% | 20%    | 10%   | 16%    | 18%    | 10%     |
| Actinobacteria         | 4%      | 3%     | 5%      | 5%     | 4%     | 4%     | 4%  | 4%     | 7%    | 5%     | 7%     | 6%      |
| Chloroflexi            | 3%      | 2%     | 3%      | 6%     | 4%     | 6%     | 4%  | 3%     | 7%    | 2%     | 3%     | 7%      |
| Spirochaetes           | 2%      | 3%     | 1%      | 3%     | 2%     | 3%     | 4%  | 2%     | 2%    | 1%     | 1%     | 1%      |
| Verrucomicrobia        | 2%      | 2%     | 1%      | 2%     | 1%     | 2%     | 2%  | 1%     | 2%    | 0.9%   | 1%     | 2%      |
| Cyanobacteria          | 2%      | 1%     | 2%      | 2%     | 2%     | 2%     | 2%  | 2%     | 3%    | 1%     | 2%     | 3%      |
| Planctomycetes         | 2%      | 1%     | 2%      | 3%     | 2%     | 3%     | 3%  | 2%     | 3%    | 2%     | 2%     | 3%      |
| Chlorobi               | 1%      | 1%     | 1%      | 1%     | 2%     | 1%     | 1%  | 1%     | 1%    | 0.9%   | 1%     | 1%      |
| Acidobacteria          | 1%      | 1%     | 1%      | 2%     | 1%     | 2%     | 1%  | 0.9%   | 2%    | 0.8%   | 1%     | 2%      |
| Thermotogae            | 1%      | 0.9%   | 1%      | 1%     | 1%     | 1%     | 1%  | 0.9%   | 1%    | 0.8%   | 1%     | 1%      |
| unclassified reads     | 0.6%    | 0.6%   | 1%      | 1%     | 2%     | 1%     | 0.9%| 1%     | 0.6%  | 1%     | 1%     | 0.7%    |
| Synergistetes          | 0.8%    | 0.6%   | 1%      | 0.7%   | 0.7%   | 0.6%   | 0.6%| 0.8%   | 0.8%  | 0.6%   | 0.8%   | 0.9%    |
| Lentisphaerae          | 0.8%    | 0.5%   | 0.3%    | 0.4%   | 0.3%   | 0.4%   | 0.4%| 0.3%   | 0.3%  | 0.2%   | 0.4%   | 0.2%    |
| Fusobacteria           | 0.6%    | 0.8%   | 0.6%    | 0.5%   | 0.6%   | 0.5%   | 0.5%| 0.7%   | 0.4%  | 0.7%   | 0.7%   | 0.4%    |
| D. Thermus             | 0.5%    | 0.4%   | 0.6%    | 0.9%   | 0.6%   | 0.8%   | 0.7%| 0.6%   | 1%    | 0.5%   | 0.6%   | 1%      |
| Aquificae              | 0.5%    | 0.4%   | 0.4%    | 0.5%   | 0.5%   | 0.5%   | 0.5%| 0.4%   | 0.4%  | 0.3%   | 0.4%   | 0.5%    |
| Tenericutes            | 0.3%    | 0.6%   | 0.3%    | 0.2%   | 0.4%   | 0.3%   | 0.2%| 0.3%   | 0.2%  | 1%     | 0.5%   | 0.3%    |
| Deferribacteres        | 0.3%    | 0.3%   | 0.3%    | 0.4%   | 0.4%   | 0.4%   | 0.3%| 0.3%   | 0.3%  | 0.2%   | 0.3%   | 0.3%    |
| Nitrospira             | 0.3%    | 0.2%   | 0.3%    | 0.4%   | 0.4%   | 0.4%   | 0.3%| 0.3%   | 0.4%  | 0.2%   | 0.3%   | 0.4%    |
| Chlamydiae             | 0.3%    | 0.2%   | 0.1%    | 0.2%   | 0.2%   | 0.1%   | 0.2%| 0.1%   | 0.2%  | 0.08%  | 0.2%   | 0.1%    |
| Fibrobacteres          | 0.2%    | 0.5%   | 0.1%    | 0.2%   | 0.1%   | 0.2%   | 0.2%| 0.2%   | 0.08% | 0.2%   | 0.2%   | 0.08%   |
| Dictyoglomi            | 0.2%    | 0.2%   | 0.2%    | 0.4%   | 0.3%   | 0.4%   | 0.3%| 0.3%   | 0.3%  | 0.2%   | 0.2%   | 0.3%    |
| Elusimicrobia          | 0.2%    | 0.2%   | 0.1%    | 0.1%   | 0.2%   | 0.1%   | 0.1%| 0.1%   | 0.1%  | 0.2%   | 0.2%   | 0.1%    |
| Chrysiogenetes         | 0.08%   | 0.06%  | 0.07%   | 0.07%  | 0.07%  | 0.09%  | 0.06%| 0.07%  | 0.09% | 0.05%   | 0.07%   | 0.09%  |
| Gemmatimonadete        | 0.06%   | 0.06%  | 0.07%   | 0.1%   | 0.1%   | 0.1%   | 0.1%| 0.07%  | 0.2%  | 0.6%   | 0.08%   | 0.2%    |
| C. Poribacteria        | 0.03%   | 0.03%  | 0.03%   | 0.05%  | 0.03%  | 0.05%  | 0.03%| 0.03%  | 0.05% | 0.02%   | 0.03%   | 0.04%   |
Among the identified species, the affiliates of *Firmicutes* were the most abundant in seven out of twelve reactor treatments and comprised 15-34% of total hits (Table 1) while the *Proteobacteria* species (Fig. 3) dominated in the other five treatments, comprising of 19-32% of total hits, (Table 1). *Proteobacteria* richness was significantly higher compared to that of *Firmicutes* species. The *Proteobacteria* communities were classified into 38 orders, which were affiliated to 7 classes (SFig. 4a). We linked their versatility with their ability to circumvent the oxidative stress in the systems. Significantly high levels of similarity of microbial composition were revealed among the treatments 1, 3 and 6, and treatments 4, 7 and 12, all located on the upper and lower right quadrant of the plot, (SFig. 4b and 5; at the class and order level). However, low β-diversities were observed for the identified *Proteobacteria* communities of reactor 4, 7, 8 and 9, (SFig. 5) at the order level. Other treatments were found to have high β-diversities of *Proteobacteria* species.

The communities’ affiliated to the class *γ-Proteobacteria* comprised 4-11% of the total identified reads. The affiliates were the fifth most abundant group in the entire microbiome and were classified into 14 orders (SFig. 7a). Among the identified orders was *Pseudomonadales*, that dominated the class, comprising of <1 to 6% of the total identified communities. This phylotype was followed by *Alteromonadales* and *Enterobacterales* species in abundance. The members of *Chromatiales*, *Vibrionales*, *Methylococcales*, *Oceanspirillales* and *Pasteurelales* respectively were also among the detected species. Other identified orders were *Xanthomonadales*, *Thiotrichales*, *Legionellales*, *Aeromonadales*, *Acidithiobacillales* and *Cardiobacterales*, all comprising <1% of the total identified communities. Unclassified *γ-Proteobacteria* reads were also identified in this study. Our results also revealed significantly high β-diversity (*P< 0.05*) of the microbial communities in the following six reactor treatments (reactor 2, 5, 6, 10, 11, 12). In contrast, the communities identified in reactors 1, 3, 4, 7, 8 and 9 were found to cluster on the upper right quadrant of the PCoA plot (SFig. 7b), indicating homogeneity and genetic composition similarity within the microbial populations.
Communities of β-Proteobacteria were detected in low abundances (2-4% of the identified microbial communities). Affiliates of Burkholderiales comprised of 1-2% of the microbial populations while the affiliates of Rhodocyclales, Neisseriales, and Nitrosomonadales species were found to comprise of <1% of the total identified populations across the treatments. Other identified β-Proteobacteria species were assigned to the order Methylphilales, Gallionellales and Hydrogenophilales which comprised of <1% of the total identified communities in the twelve treatments. High sample to sample variation was also observed for the β-Proteobacteria species (SFig. 8b).

The affiliates of α-Proteobacteria were revealed to comprise of 3-8% of the total microbial communities, and were classified into 7 orders (SFig. 9a). The Rhizobiales dominated this group, comprising of 1-2% of the total microbial population in the treatments. The relative abundances of Rhodobacterales, Rhodospirillales, and Sphingomonadales followed (representing <1% of the total population). Other α-Proteobacteria orders identified included Caulobacterales, Rickettsiales and Parvularculales. We also found significantly high dissimilarities among the α-Proteobacteria within the studied treatments. Only the α-Proteobacteria communities of reactor 3 and 11 were revealed to express low β-diversity (SFig. 9b), a trait that we postulate was attributed by similar reactor operating conditions rather than the observed environmental conditions (Table 1 and 2).

The ε and Z-Proteobacteria were among the least abundant Proteobacteria phylotypes in our study and were found to comprise ≤1% of the total identified microbial populations in the treatments. The ε-Proteobacteria communities were classified into the orders Campylobacterales and Nautiliales (SFig. 10a), both comprising <1% of the total identified populations. We found similarities between the ε-Proteobacteria communities of reactor 4 and 8 (clustered) and those of reactor 5 and 9 (clustered partially) revealing low β-diversity (SFig. 10b). All the identified Z-Proteobacteria were from the order Marinobacterales, and accounted for <1% of the total microbial populations in the studied treatments.

The Firmicutes communities were the second most diverse community identified in this study. They comprised of 15-34% of the total microbial communities identified. The community was classified into 8 orders and 4 classes (SFig. 11a and 12a). The Firmicutes in the three treatments of reactor 2, 10 and 12 revealed significant genetic similarity and were revealed to cluster in the lower left quadrant of the PCoA plot (SFig. 11b). The Firmicutes in reactor 1 and 7 were found to cluster partially on the upper right quadrant of the PCoA plot (significant threshold; P≤0.05; SFig. 11b), indicating low β-diversity. At the order level, the identified Firmicutes of reactor 2 and 10; 3 and 6; 7 and 12; and reactor 1 and 9 also exhibited low β-diversities (SFig. 12b). Other local communities were revealed to express high β-diversities at class and order levels.

The Clostridia was revealed to be one of the most abundant Firmicutes phylotype, comprising 8-22% of the identified total microbial communities. It was classified into 4 orders (SFig. 13a). The order Clostridiales dominanted the phylum and constituted 8-22% of the total communities. Thermoanaerobiales species constituted 2-3% of the total identified communities. Other identified Clostridia species were the affiliates of Halanaerobiales and Natranaerobiales, all represented <1% of the total microbial populations, except for the Halanaerobiales communities in reactor 11 that comprised of 2% of the total identified microbial populations. Among the twelve biogas reactor treatments, three treatments (i.e reactor 7, 11 and 12) clustered on the lower left quadrant of the PCoA plot, (SFig. 13b), indicating high community similarities. Communities of reactor 5 and 9 were found to cluster partially on the upper left quadrand of the PCoA plot, revealing significantly low β-diversity.
The other seven treatments had high β-diversity, implying metabolic and genetic plasticity among the identified Clostridia species. The class Bacilli comprised of 4-9% of the total identified communities and was the fourth most abundant Clostridia. The Bacilli were classified into Bacillales and Lactobacillales (SFig. 14a), which comprised 3-4% and 1-4% of the total identified microbiome communities, respectively. Among the Bacilli communities, those of reactor 5 and 8 were found to exhibit significant dissimilarities, while the others exhibited a low β-diversity (SFig. 14b). Other detected low abundance Firmicutes were Erysipelotrichales (Erysipelotrichi) and Selenomonadales (Negativicutes) which comprised of <1% of the total identified communities. The Erysipelotrichales species in reactor 5 and 9 were however found to comprise of 1% of the total identified microbial populations.

Bacteroidetes were the third most abundant phylotype comprising of 10-22% of the total identified microbial populations. These were classified into 4 classes and orders (SFig. 15a and SFig. 16a). Among them, were the Bacteroidia that comprised of 5-12% of the identified microbial communities. All the Bacteroidia belonged to the order Bacteroidales. Other members of Bacteroidetes were Flavobacteria that belong to the order Flavobacterales. The communities of Cytophagia were all assigned to Cytophagales species and they comprised of 1-2% of the total identified microbial communities. The other Bacteroidetes were Sphingobacteria, mainly of the order Sphingobacteriales, which comprised of 2-3% of the total identified microbial communities. We observed partial clustering of the Bacteroidetes’ communities particularly in reactor 2 and 10 and in reactor 8 and 12, which indicated low β-diversity between the respective communities (SFig. 15b and 16b). The Bacteroidetes communities of reactor 4 and 9 also clustered on the lower left quadrant of the PCoA plot, (SFig. 15b and 16b), implying significant similarities of the Bacteroidetes in the two treatments.

We identified 6 orders of Actinobacteria species (SFig. 17a) which comprised of 3-7% of the identified microbial reactor communities. The Actinomycetales¹ were among the identified affiliates, comprising of 2-6% of the total identified communities. The species of order Coriobacteriales and Bifidobacteriales accounted for <1% of the total identified microbial composition. Other less abundant Actinobacteria species were assigned to the order Rubrobacteriales, Solirubrobacteriales and Acidimicrobiales. Out of the twelve biogas reactor treatments, the microbial communities of the seven treatments (reactor 1, 3 and 12; reactor 5 and 11 and reactor 7 and 8; SFig. 17b) were found to exhibit significantly (P<0.05) low β-diversities. Only Actinobacteria species of the remaining three out of the twelve treatments were found to exhibit high β-diversities.

The Chloroflexi were categorized into 4 classes (SFig. 18a) and 5 orders (SFig. 19a) and comprised of 2-7% of the total microbial communities identified (Table 3). The communities belonged to Chloroflexi (1-4% of the identified communities) and Dehalococcoidetes (≤1-3% of the identified communities) as the most abundant phylotypes. Although all the Dehalococcoidetes reads were unclassified at a lower level, we identified Chloroflexales and Herpetosiphonales as the affiliates of the class Chloroflexi (SFig. 20). The two orders comprised of 1-2% and <1% of the total identified microorganisms, respectively. Other identified classes of the phylum were Thermomicrobia and Ktedonobacteria; all of which comprised of <1% of the total identified communities. The Thermomicrobia were classified into orders Sphaerobacterales and Thermomicrobiales (SFig. 21a). All the identified Ktedonobacteria belonged to the order Ktedonocacterales (also representing <1% of the identified communities). We identified two treatments (reactor 1 and 7) that revealed low β-diversities of their local Chloroflexi communities (SFig. 18b). Four treatments (reactor 4, 5,
out of the remaining ten were revealed to also cluster on the lower right quadrant of the PCoA plot (SFig. 18b and 19b).

Similarly, the Chloroflexi communities of reactors 1 and 5 and those of reactors 3 and 7 were found to cluster on the upper right and lower right quadrants of the PCoA plot (SFig. 20b). The Chloroflexi class communities of reactor 1 and 10 were found to cluster partially with the cluster of microbes from reactor 3 and 7. We also identified partial clustering of Chloroflexi from reactor 8, 11 and 12 indicating low β-diversities among the communities in these reactors (SFig. 20b). The Chloroflexi from the rest of the treatments revealed high β-diversities. In contrast, the Thermomicrobia communities were significantly dissimilar among the treatments as indicated in SFig. 21b.

The Cyanobacteria comprised of 1-3% of the total microbial composition and majority in this group were unclassified. Only <1% of the total population were assigned to Gloeobacteria (SFig. 22a). Unlike other phylotypes, only the Cyanobacteria communities of reactors 3, 7 and 10 were found to exhibit significantly high β-diversity ($P \leq 0.05$), (SFig. 22b) while the communities in the other treatments were found to express partial similarities between or among themselves. The Gloeobacteria communities were all classified as Gloeobacteriales order. Other identified communities (unclassified reads) included members of the following orders; Chroococcales, Nostocales, Oscillatoriales and Prochlorales (SFig. 23a). Overall, a majority of the local Cyanobacteria communities were found to reveal low β-diversities. The Cyanobacteria communities in reactors 4, 6, 9 and 12 however exhibited high β-diversities (SFig. 23b). Similarly the identified four phylotypes at the order level (derived from unclassified reads, SFig. 24a) were also revealed to express high local compositional dissimilarities among the twelve biogas treatments (SFig. 24b).

We also identified Acidobacteria which comprised of 1-2% of the total identified microbial communities. Majority of these populations were from the order Acidobacteriales (comprising <1-2% of the total identified microbial communities). The remaining species were affiliated to Solibacteres, particularly of the order Solibacterales (<1% of the total identified microbial communities) (SFig. 25a and 26a). The Acidobacteria communities of five (reactors 1, 2, 3, 8, 12) out of twelve treatments were revealed to exhibit low β-diversities. Acidobacteria communities of reactor 2 and 8 clustered on the upper left quadrant of the PCoA plot (at order level; SFig. 25b and 26b), indicating similarity between the two local communities. Deinococcus-Thermus, another phylum, had the orders Deinococcales and Thermales species (SFig. 27a), the two belonging to the class Deinococci. The relative abundances of the two orders were ≤1% of the total identified microbial populations. We observed low β-diversities between the Deinococci species of reactor 11 and 12 while the diversity of its communities in reactors 3 and 6 and those of reactors 7 and 9 were found to be significantly similar (SFig. 27b). All the other local Deinococcus-Thermus communities expressed high β-diversities.

The identified Verrucomicrobia communities were classified into 3 classes that comprised 1-2% of the total identified microbial communities. Among the identified classes were Opitutae, Verrucomicrobiae and Spartobacteria (SFig. 28a). The affiliates of Verrucomicrobiales were the most abundant followed by those of the Puniceicoccales and methylacidiphilales species; all representing <1% of the total identified communities (SFig. 29b). Notably, we found out that all the Opitutae and Spartobacteria reads were unclassified at a lower taxonomic level. Only the Verrucomicrobia communities of reactors 3 and 7 were found to cluster partially on the lower right quadrant of the PCoA plot, (at the class level, SFig. 28b). In contrast, several local communities including the ones identified in reactors 1
and 10, and reactors 3, 6, 7 and 11 were found to cluster partially (indicative of low β-diversities) within the plots (at order level; SFig. 29b). However, their communities in reactor 5 and 8 were found to cluster significantly on the upper right quadrant of the PCoA plot (SFig. 29b), implying high similarities between the Verrucomicrobia species in the two treatments. Other treatments were found to reveal high β-diversities. Among the identified Tenericute’s, were Acholeplasmatales, Mycoplasmatales and Entomoplasmatales, all comprising of <1% of the total identified microbial communities and belonged to the class Mollicutes (SFig. 30a). Interesting, member species of this class were absolutely absent in reactor 5 only, contrary to other phyla species, probably due to natural selection pressure. Our data also revealed significantly high similarities of the identified Tenericute’s species inhabiting reactors 4 and 9 and reactors 6 and 8 (SFig. 30b). The local communities of the other treatments were observed to be significantly dissimilar.

Our results also identified Planctomycetacia affiliates which comprised 1-3% of the total identified microbial communities. These affiliates were all assigned to the order Planctomycetaceae (Planctomycetacia). Chlorobia were identified in our study (comprising of 1-2% of the total identified communities). These belonged to the order Chlorobiales. Spirochaetes were also identified (1-4% of the total identified microbial composition). All the Spirochaetes identified in this study belonged to the order Spirochaetales. Thermotogae of the order Thermotogales were also identified in the study. Other less abundant communities identified were Synergistales (Synergistata, Synergistia), Lentisphaerales (Lentisphaerae, Lentisphaeraceae), Fuscobacteriales (FusoBacteria, FusoBacteria), and Aquificales (Aquificae, Aquificae), all of which comprised of <1% of the total identified communities. DeferriBacteriales, Chlamydiales, Chrysiogenales, Nitrospirales, Fibrobacteriales, Elusimicrobiales, Dictyoglomales, Gemmatimonadales were other groups of microbes identified in this study. The Candidatus poribacteria affiliates were not accurately classified at the lower taxonomic ranks (Table 3).

The identified Archaeal biomes
We analyzed the identified Archaea domain and established that their species comprised of 6-12% of the total microbial populations. Out of the identified communities, we obtained 5 phyla (SFig. 31a), 9 classes (SFig. 32a) and 16 orders (SFig. 33a). Among the communities, those identified in reactor 4 and 12 were found to cluster on the lower left quadrant of the PCoA plot, indicating their high similarity. We also revealed low β-diversities of the identified Archaeal communities, of reactor 3 and 6 and also those detected in reactor 7 and 9 (Phylum level; SFig. 31b). However, at a lower taxonomic level, only those identified in reactor 3 and 7 were observed to have low β-diversity (Class and order level; SFig. 32b and 33b). All the other reactor treatments were observed to contain highly dissimilar communities of Archaea at different taxonomic ranks.

A majority of Archaea species identified in this study were classified as Euryarchaeota and comprised of 5-11% of the total microbial communities. The Euryarchaeota communities were classified into 8 classes (SFig. 34a) and 10 orders (SFig. 35a). Interestingly, we observed significant dissimilarities (high β-diversity) of the Euryarchaeota among the reactor treatments at the two taxonomic levels of class and order (SFig. 34b and 35b). The class Methanomicrobia was the most abundant Euryarchaeota phylotype and comprised of 4-9% of the total identified microbial populations. Further, we established that the majority of the members of the class were affiliates of Methanomicrobiales order (2-4% of the identified total communities), followed by those assigned to Methanosarcinales and Methanocellales (SFig. 36a), which comprised 1-3% and ≤1% of the identified total microbial communities, respectively. The local composition of Methanomicrobia species were highly dissimilar.
among the reactor treatments except for those from reactor 1 and 7 that were found to cluster. The communities of the two treatments revealed significant similarities (significant threshold, \( P \leq 0.05 \); SFig. 36b). Other identified methanogenic communities were assigned to *Methanobacteria*, the class *Methanococcales* belonging to *Methanococci* class; and *Methanopyrales* the affiliates of *Methanopyri* phylotype. All these species were revealed to comprise of <1% of the detected microbial communities.

Other non-methanogenic *Euryarchaeotes* identified in our study were the affiliates of *Thermococcales* (*Thermococci*), *Archaeoglobales* (*Archaeoglobus*), *Halobacteria* (*Halobacteria*), and *Thermoplasmatales* (*Thermoplasmata*). Among these phylotypes of *Euryarchaeotes*, affiliates of *Thermococcales’* were dominant while the *Thermoplasmatales* were the least abundant. *Crenarchaeota* and *Thaumarchaeota* phyla were among the other identified non-methanogenic species. Between the two phyla, *Crenarchaeota* were the most diverse and were classified into 4 orders (SFig. 37a). Within the identified *Crenarchaeota*, the following orders were identified; *Desulfurococcales*, *Thermoproteales*, *Sulfolobales* and *Acidilobales* (in order of abundance). All these orders are members of *Thermoprotei* class which comprised of <1% of the identified microbial communities. We noted that half of the studied treatments (reactor 6 and 9, reactor 7 and 10 and reactor 2 and 11) had *Crenarchaeota* communities that had low β-diversities, while the rest of the communities, in the other six treatments had significantly high β-diversities (significant threshold, \( P \leq 0.05 \); SFig. 37b). *Thaumarchaeota*, another non-methanogenic phylum was classified into *Nitrosopumilales* and *Cenarchaeales* orders (SFig. 38a) both of which exhibited high β-diversities (SFig. 38b). Other identified phyla were *Korarchaeota* and *Nanoarchaeota*. The classification of these two phyla at the lower taxonomic ranks was inaccurate and the sequence reads from the two were therefore treated as unclassified.

**The identified fungal biomes**

Our study also identified the presence of *fungal* species. The fungal communities were classified into 5 phyla (SFig. 39a), 14 classes (SFig. 40a) and 23 orders (SFig. 41a). Their relative abundances were however low and they comprised of <1% of the total microbial populations among the reactor treatments. Significantly high β-diversities were noted for the identified local communities of *fungi* at the phyla and class levels (SFig. 39b; SFig. 40b). However, at the lower taxonomic levels, low β-diversities were a common feature particularly for the communities identified in reactor 2 and 10, reactor 3 and 6, reactor 4 and 9, and reactor 8 and 11 (SFig. 41b).

The affiliates of *Ascomycota* were the most abundant guild (<1 of the total communities) within the *Eukaryotes* and were classified into 7 *Ascomycota* classes (SFig. 42a) that included *Eurotiomycetes*, *Saccharomycetes* and *Sordariomycetes*, respectively. Other identified classes were *Schizosaccharomycetes*, *Dothideomycetes*, *Leotiomycetes* and *Pezizomycetes* all of which comprised <1% of the total microbial communities in the respective treatments. We also further classified the identified *Ascomycota* species into 11 orders (SFig. 43a). Interestingly we observed significantly low β-diversity between *Ascomycota* communities in reactor 7 and 8. Other identified *Ascomycota* communities were observed to exhibit high β-diversities at the lower taxonomic levels (class and order level; SFig. 42b and 43b). Among the classes, *Sordariomycetes* were the richest within the phylum and their species were classified into 4 orders; the *Hypocreales*, *Sordariales*, *Magnaporthales* and *Phyllachorales* (SFig. 44a). The *Sordariomycetes’* communities were found to exhibit significantly high β-diversities (Statistical threshold, \( P \leq 0.05 \)) (SFig. 44b) among the studied treatments. Despite their higher abundance compared to *Sordariomycetes’* phylotypes, the *Eurotiomycetes’* another class of *Ascomycota* were classified into two orders; *Eurotiales* and *Onygenales*.
The two Eurotiomycetes orders were shown to comprise <1% of the identified total microbial populations. Like other fungal phylotypes, only the Eurotiomycetes’ communities of reactor 1 and 3 were found to cluster on the PCoA plot, which was indicative of high community similarity. The other reactor communities were revealed to have significantly dissimilar Eurotiomycetes populations (significant threshold, $P \leq 0.05$; SFig. 45b). Other identified fungal species were the affiliates of Saccharomycetales, Schizosaccharomycetales, Pleosporales, Helotiales and Pezizales phylotypes, all comprising <1% of the total microbial communities.

Among the identified fungi, Basidiomycota species were the second most abundant (<1% of the total communities). The species were classified into 4 classes and 5 orders (SFig. 46a and 47a, respectively). Among the classes, Agaricomycetes dominated the treatments, followed by Tremellomycetes, Ustilaginomycetes, and Exobasidiomycetes species (SFig. 46a) respectively. The affiliates of the class Agaricomycetes were further classified as Agaricules and Polyporales species (SFig. 48a). Other Basidiomycota species were affiliated to the following orders; Tremellales (Tremellomycetes), Ustilaginales (Ustilaginomycetes) and Malasseziales (Exobasidiomycetes). We observed two treatments’ (reactor 11 and 12) that contained Agaricomycete’s communities with low β-diversities (SFig. 48b) while the local Basidiomycota communities in the other reactor treatments had significantly high β-diversities (SFig. 46b and 47b).

The Chytridiomycota were among the identified rare species. This was classified into Chytridiomycetes (more abundant) and Monoblepharidomycetes (SFig. 49). Among the Chytridiomycota were Spizellomycetales (identified in reactor 9 and 10), Rhizophydiales and Cladochytriales, the latter two detected only in reactor 9 (SFig. 50). All the affiliates of Monoblepharidomycetes were classified as Monoblepharidales species and were also detected in reactor 9 only. Other rare phylotypes identified in our study include Blastocladiales, (Blastocladiomycetes, Blastocladiomycota, detected in reactor 9), Entomophthorales, Mortierellales, and Mucorales, all derived from unclassified fungal reads (also identified in reactor 9, 10 and 11 respectively, SFig. 51). Microsporidia species were not accurately classified at the lower taxonomic levels and were regarded as unclassified though they comprised <1% of the total microbial populations.

The microbiomes’ α-diversities and biogas production

Since environmental and technical variables can outweigh the biological variations when handling metagenomic datasets, prior to analysis, we asked ourselves, whether we had attained the maximum microbial richness and if our sampling was a representative of this maximum. To answer these questions, we conducted a rarefaction analysis through statistical resampling and plotted the total number of distinct species annotations against the total number of sequences sampled (Fig. 4). As a result, a steep slope, that leveled off towards an asymptote was obtained for each sample and later combined (Fig. 4). This finding indicated that a reasonable number of individual species were sampled and more intensive sampling of the same treatments would yield probably few additional species, if not none. Further the α-diversity indices revealed 947-1116 genera across the treatments, confirming our maximum sampling effort was attained for further diversity studies. The study also revealed variabilities of reactor performance i.e biogas productivity (Table 3) among treatments, the phenotype that we linked to the observed α-diversity variability, among the treatments.
The microbiomes’ β-diversities and biogas production

To understand the influence of environmental variations and other variables on the local microbial communities and thus biogas production, we conducted comparative community diversity studies on the twelve biogas reactor treatments. As a result, we revealed high β-diversity among the local communities in the majority of the treatments (at the domain level). The only observed exceptions were the local communities of reactor 4 and 11 that clustered on the lower left quadrant of the plot (SFig. 52a), revealing significant similarity between the organisms in the two treatments. Similarly, at a lower taxa level (phylum, class and order), only the communities of reactor 2 and 10 were found to be similar, and were located in either the lower right or upper left quadrants of the PCoA plot (SFig. 52b and 52c). We also observed low β-diversity between the communities of reactor 4 and 9 at the lower taxonomic levels (phylum, class and order) with species clustering partially on the lower left and upper right quadrants of the PCoA (SFig. 52b and 52c), respectively. Other treatments that revealed low β-diversity between their communities were reactor 8 and 11 and reactor 3 and 6. Communities from these reactors were found to cluster partially in the lower right and lower left quadrants of the plot (Fig. 5). Similarly, we also noted significant variability of biogas production among the treatments (Table 3), an indication of the observed β-diversities at different taxonomic levels and within the taxa (stated above).
**Fig. 5:** The PCoA analysis revealing β-diversity of the twelve treatments at the order level: The reads compositions in reactor 2 and 10 were partially similar, located in the lower left quadrant, those in reactor 8 and 11 located in the upper right quadrant, reactor 4 and 9 in the lower right quadrant, while those in reactor 3 and 6 were located in the lower left quadrant of the plot.

**Discussion**

**Annotations against the 16 biological databases**

The aim of our study was to use next generation sequencing technologies and different of bioinformatics tools to investigate the biogas microbiomes from environmental samples obtained from a wide geographic coverage. We combined marker genes and other genes on the assembled scaffolds to generate maximum taxonomic assignments. Among the sixteen utilized databases, SEED subsystems, yielded the highest reads hits annotations, the findings that corroborated with the previous reports [22]. Unlike other databases, SEED integrated several evidences that provided annotation benefits over the other 15 utilized databases. PATRIC database had fewer annotations compared to the IMG database, due to its high speciality of storing only pathosystems information. The annotation of low protein sequences by the SWISS-PROT database was due to its manual curation and annotation systems. Contrasting the findings were the annotations of TrEMBL and Refseq databases that both utilized automated schemes, which superseded the annotation of SWISS-PROT database. However, in spite of its low resolution, majorly due to lack of regular updates, the outputs of SWISS-PROT are more accurate compared to the automated databases. Furthermore, orthologous groups databases such as COG, eggNOG etc were utilized to distinguish orthologs from paralogs. In this case, we coupled gene classification schemes to the taxonomic assignments in order to increase detection resolution. Among these five databases (Fig. 1), eggNOG had the highest hits annotation due to its high phylogenetic resolution, automated annotation scheme, and coverage of more genes and genomes than its counterpart databases. In contrast, COG and NOG databases were among the databases that had the lowest hits annotations because of their manual curation and annotation schemes and fewer regular updates due to manual labor requirements. KEGG database was used to reconstruct metabolic pathways of the identified taxonomic assignements. Other utilized nucleotide databases were Genbank that ensured uniform and comprehensive annotation against the worldwide nucleotide sequences, Ribosomal Database Project (RDP) database that covered small rRNA gene subunit (SSU) of Archaea and Bacteria domains, LSU and SSU databases that archived large and small rRNA subunit of Archaea, Bacteria and Eukarya sequences respectively. The latter two; provided a complete species detection advantage over the RDP database. The two databases archived nucleotide sequences together with their postulated secondary structures. The SSU database stored more community sequences than its counterpart; the LSU database and it used prokaryotic and eukaryotic models for secondary structure alignments. The extra feature of the SSU database explained why the database, had the highest annotation resolution compared to other amplicon sequences databases. However, its worth-noting that the three rRNA databases archived complete or near-complete rRNA sequences, although partial fragments of 70% of the estimated chain length of the combined molecule were also included. These characteristics of the amplicon sequence databases limited the annotation of our Miseq generated short reads. The ENVO ontologies analysis, which revealed high resolution of large lake biomes, suggested that majority of the cultured and sequenced species were derived from large water masses.
Annotations against the archived taxonomic profiles

Our taxonomic annotations approach generated more detection, compared to the other annotations conducted elsewhere [23]. While other studies revealed 0.042% to 0.132% rRNA genes, we revealed 280 (0.00%) to 1.032 (0.00%) rRNA genes. These findings implied that the proportion of the rRNA genes to the total shotgun generated genes were negligible. The observations were highly attributed to our sampling and sequencing approaches as previously explained by Knight et al. [24]. Furthermore, while most of the studies used 19-35% of the total generated sequences, [23], our approach annotated and used 39.75-46.70% of our total generated sequences. All these findings indicate that more than 53% of the biogas generating populations is still unknown. Briefly, our annotations were assigned to Bacteria, Archaea, Eukaryotes and Viruses, which corroborated the findings of other studies [25]. Among the domains, Bacterial species dominated in abundance, comprising 87-93% of the total microbial populations, followed by Archaea and Eukaryotes, while Viruses comprised the least abundant population. The observed high Bacterial abundance, suggest their crucial metabolic roles in biomass conversion and other reactions within the reactor systems [25]. As expected, we found low abundance among the Eukaryotes (fungi) which was indicative of the presence of low recalcitrant biomasses concentration. Generally, the fungal species are involved in several biochemical reactions of biogas production including hydrolysis, fermentation and syntrophic reactions. The identified affiliates of Archaea were majorly consumers of smaller substrates that were generated by the Bacterial and fungal species. The Archaea species are able to use different methanogenic routes to convert the substrates into CH₄(g). Nevertheless, the main roles of the identified and less abundant group of the virus were unclear, though the species could have been active in degrading other microbial cells in the AD systems. However, their detection was largely attributed to the viral pathogenicity on both Bacteria and Archaea species. High β-diversity among the local communities (at the domain level) was an indication of environmental influences at the different ecological location.

The identified Bacterial biomes

More than 90 representative orders of the Bacteria domain were revealed in our datasets. The first step of the AD process involves hydrolytic reactions that convert large macromolecules into smaller substrates [26]. Several Bacterial communities are capable of hydrolysis, even on high lignocellulosic plant biomasses. The identified species that participate in these reactions were annotated to the Firmicutes, Bacteroidetes, Chloroflexi to mention but a few and their species are known to harbor cellulosomes, that secrete extracellular enzymes, to degrade the substrate. Previous reports have indicated that species belonging to Bacteroidales, Bifidobacteriales, Clostridiales Lactobacillales, Thermotogales and many other identified Bacterial species catalyze the first step of AD [26]. The step involves rate limiting reactions, largely regulated by the substrate composition. The nature of the chemical composition of the substrate, generally depends on the environmental variables (Table 2 and 3), which influence microbial species’ abundances and β-diversity differently. All these affects reactor stability and performance. As expected, the class Clostridia, that comprised Clostridiales, Thermoanaerobacteriales and Halanaerobiales orders, was the most dominant among the identified phylotypes. The guild converts large macromolecules into organic acids, alcohols, CO₂(g) and H₂(g) [27-33]. In this case, the species use a substrate-level phosphorylation mechanism to conserve the energy. Acetogenic and homoacetogenic species belonging to the same guild such as Moorella thermoacetica and Thermoanaerobacter kivui are able to use the reductive acetyl-CoA pathway to reduce CO₂(g) into CO₂(g), concomitantly generating reducing equivalents [34, 35] and vice versa. As a consequence, the acetogens and homoacetogens use proton or sodium dependent mechanisms to conserve their energy [35].
Above all, the *Clostridia* guilds are able to catalyze other biochemical reactions such as reductive dechlorination of various compounds that leads to production of sulfide from either thiosulfate or sulfate. The reactions consume pyruvate and lactate as the carbon source. In this case, the member species are able to use [NiFe]-hydrogenases to generate H\textsubscript{2(g)} [33]. In our systems, we also identified *Halanaerobiales* species that have a capacity to tolerate high salt concentrations [36]. Generally the reactor treatments comprised high salts concentrations and the *Halanaerobiales* must have regulated the toxic cations from the ionization of the salts. All these roles of *Clostridia* suggest their metabolic versatility, and these features were used to explain the observed low richness of *Firmicutes* compared to *Proteobacteria* species. Similarly, the observed high β-diversities among the treatments were linked to the versatile roles of *Clostridia*. However, the observed low β-diversities (SFig. 13b) of communities in reactor 7, 11 and 12 are probably attributed to similar treatment perturbations and the identified environmental variables (Table 2 and 3).

The *Bacillales* and *Lactobacillales* species identified in our study, all belonged to the class *Bacilli*. These species are able to use mixed fermentation routes to metabolize carbohydrates and other biological macromolecules [37, 38]. Formate dehydrogenase complex activities are the main mechanisms of energy conservation [38] in these microbial guilds. Generally, *Bacilli* co-exist in a symbiotic relationship with *Clostridia* species in lactate and acetate metabolism [39], a mechanism that is termed as lactate cross-feeding. High β-diversities among the observed treatments (SFig. 14b) were solely attributed to cow-dung’s substrate composition and environmental variables (Table 2 and 3). However, the low β-diversities observed between or among the treatments (SFig. 14b) that formed the clusters were largely attributed to similarities of reactor’s operation conditions. All these implied that the respective organisms were in the same succession stage at the time of sampling. The *Erysipelotrichi* another identified *Firmicutes* class, utilize the syntrophic acetogenesis route to metabolize lipids into acetate, H\textsubscript{2(g)} and CO\textsubscript{2(g)} [23]. Its species, together with other *Firmicutes* species like *Smithella propionica* use a non-randomizing pathway [40] to metabolize propionate into acetate. Among other identified *Firmicutes*, were *Syntrophomonas wolfei* (*Clostridiales, Clostridia*) that use the β-oxidation pathway to oxidize C\textsubscript{4} to C\textsubscript{7} fatty acids in order to produce either acetate and H\textsubscript{2(g)} or propionate, acetate and H\textsubscript{2(g)} [41, 42]. Notably, these syntrophic species are consumers of 30% of the generated electrons, which bypasses the rate limiting steps of the volatile fatty acids accumulation. All these, contributes to reactor stabilities. In general, we associated the observed low *Firmicutes* abundances and diversities (richness, SFig. 11a and 12a) with genetic versatility that conferred a capacity to encode an assortment of proteins; which metabolize a variety of biomasses into biogas. Furthermore, the observed high β-diversities (SFig. 11b and 12b) of the *Firmicutes* support our concepts on the genetic divergence among the species and we postulate that the trait was mainly influenced by the sizes of the reactors and the variability of the environmental conditions (Table 2 and 3).

The *Synergistete’s* species identified in this study have been established to ferment amino acids into lactate, acetate, H\textsubscript{2(g)} and CO\textsubscript{2(g)} [43, 44]; though they have no ability to metabolize carbohydrates and fatty acids. One such identified example is *Dethiobullovibrio peptidovorans*. The affiliates of genus *Anaerobaculum* which belong to *Synergistetes* catabolize peptides into acetate and short-chain fatty acids while their close relatives, *Aminobacterium*, are capable of fermenting amino acids into acetate, propionate and H\textsubscript{2(g)} [43, 44]. Other amino acid metabolizers identified in this study were the affiliates of *Fusobacteria* [45] which are able to catabolize substrates into acetate, propionate, and butyrates, concomitantly releasing H\textsubscript{2(g)}, CO\textsubscript{2(g)} and NH\textsubscript{3(g)}. The member species of the two phyla are known to use the substrate-level phosphorylation to conserve their energy [46] from amino
acids metabolism. Our findings contradict the metabolism of the amino acid by the Clostridia species, whose energy conservation mechanisms remain unclear and have been postulated to occur through chemiosmosis [30]. However, the two phyla described here have been found to generate energy through syntrophic association with methanogens; otherwise the generated protons would increase the systems’ alkalinity and lead to inhibited steady-state of the biochemical reactions. Based on the low abundance of the two phyla and richness among the treatments, we postulate low amino acid affinities among their identified species, which would suggest another selection pressure in the AD systems.

The identified affiliates of Bacteroidales, Flavobacterales, Cytophagales and Sphingobacterales, all belong to Bacteroidetes, and are majorly known to ferment carbohydrates and proteins into mixed products (acetate, lactate, ethanol, and volatile organic acids), concomitantly releasing H\textsubscript{2}(g) [47]. Moreover, their species are also capable of metabolizing aliphatic, aromatic and chloronated compounds [48] and were observed to co-exist with methanogens possibly to increase energy extraction from indigestible plant materials. The member species were highly varied among the treatments and we hypothesize that the cattle breeds and the operating conditions in the reactors were the key drivers of the observed β-diversities. Interestingly, we also observed consistency in the relative abundance (2% of total reads) of the Cytophagae’s guild among the different treatments, which may be attributable to the long periods in conditions of low nutrient limitation [49]. Although, the metabolic activity of the identified Chlorobi overlap with that of Bacteroidetes [50], their species use reverse tri-carboxylic acid (rTCA) cycle to fix CO\textsubscript{2}(g) and N\textsubscript{2}(g) in anaerobic conditions. We hypothesize that these rTCA reactions may have been utilized by the Chlorobi to oxidize reduced sulfur compounds [51, 52]. In our results, the Bacteroidetes species however revealed a higher selection advantage over Chlorobi due to the observed low Chlorobi abundances and richness (i.e with only a single order, the Chlorobale).

The identified Proteobacteria guilds consisted of metabolically versatile species and were revealed to comprise the highest number of species (richness) compared to other identified phyla in the treatments. The guilds comprised of species with capacity to metabolize proteins [53] carbohydrates and lipids [54] and we associated their low β-diversity (SFig. 4b and 5b) with low cow-dug’s substrate composition heterogeneity and differences in environmental conditions (Table 2 and 3). The γ-Proteobacteria was revealed in our study as the most diverse class of the Proteobacteria and the class comprised of 14 orders (SFig. 7). The Enterobacteriales included Escherichia coli, which are facultative anaerobes able to ferment carbohydrates into lactate, succinate, ethanol, acetate, H\textsubscript{2}(g) and CO\textsubscript{2}(g) [55]. Unlike other syntrophic species, the affiliates of Enterobacteriales have been found to use the methylcitrate cycle [56] to metabolize propionate into pyruvate, a precursor for butanol, isopropanol and other mixed acid fermentation products. Burkholderiales, another γ-Proteobacteria phylotype identified in this study, produces succinate from carbohydrates concomitantly oxidizing H\textsubscript{2}(g), S\textsuperscript{0} and Fe\textsuperscript{0} into their respective oxidized forms to conserve energy [57, 58]. We suggest that the high abundance and β-diversities of the γ-Proteobacteria communities (SFig. 7, Table 3) could be due to low environmental temperature.

In contrast, the affiliates of genus Sutterella, among other species of β-Proteobacteria identified in this study are known saccharolytic and nitrate reducers [58]. The species also consume propionate, butyrate and acetate syntrophically [54, 58]. We suggest that the high abundance and diversity of β-Proteobacteria (SFig. 8) in the treatments was favoured by the agroecological zones and prevailing environmental temperatures (Table 2 and 3) of the reactors. The Caulobacterales, an order of α-Proteobacteria were among the identified low abundant species in our study. These species have previously been revealed to survive in low
nutrient environments [59]. *Rhodospirillales* (α-Proteobacteria), such as *Rhodospirillum rubrum*, have been previously determined to participate in homoacetogenesis and convert \( \text{CO}_2(g) \) and formate into acetyl-CoA, \( \text{CO}_2(g) \) and \( \text{H}_2(g) \) [60]. The produced acetyl-CoA is further converted into acetate, a precursor for methanogenesis. Like members of β-Proteobacteria, the α-Proteobacteria species identified in this study were also influenced by the agroecological zones (Table 2 and 3) and we postulate that the cow-dug’s substrate composition was the main cause of the observed β-diversities and variations (SFig. 9) among the treatments. The *Campylobacterales* and *Nautiliales* species which are affiliates of ε-Proteobacteria were also identified in our study. These species metabolize proteins in the AD systems and have also been revealed to express an additional role of antibiotic resistance in biofilms and granules [54].

Among the identified δ-Proteobacteria species, *Pseudomonadales*, the gemmules producing anaerobes [61] were the most abundant and have been found to co-exist with algae [61]. The reported syntrophic traits of the *Pseudomonadales* order were revealed to enable the species to withstand unfavourable environmental conditions, which could have offered the species the selection pressure advantage over other δ-Proteobacteria species. Generally, the δ-Proteobacteria species are syntrophic oxidizers [62-64], with the ability to convert volatile fatty acids into acetate, formate, \( \text{CO}_2(g) \) and \( \text{H}_2(g) \). The *Syntrophobacterales* species identified in this study e.g *Syntrophus aciditrophicus* are able to use the randomized methyl-malonyl-CoA pathway to oxidize butyrate, propionate and long-chain fatty acids into acetate, formate, \( \text{CO}_2(g) \) and \( \text{H}_2(g) \) syntropically [62, 63] while *Syntrophobacter wolinii* [64] another member species, is known to use the β-oxidation pathway to metabolize propionate into acetate. Other species belonging to the genus *Syntrophaceticus* identified in this study have been reported to oxidize the produced acetate into \( \text{CO}_2(g) \) and \( \text{H}_2(g) \) [65]. Interestingly, for these species to effeciently oxidize the respective macromolecules, they have been revealed to use an electron transport mechanism that creates a positive redox potential by reversing electrons that were generated in syntrophic reactions. In contrast, the identified affiliates of *Desulfarculales*’s, *Desulfofbracteale’s* and *Desulfovibronale’s* phylotypes are sulfate reducers, although in the absence of sulfate (electron acceptor), the species are able to oxidize organic compounds, (e.g lactate, acetate and ethanol) into \( \text{CO}_2(g) \) and \( \text{H}_2(g) \) syntrophically [66]. The *Desulfoomonadale’s* species have been revealed to use \( \text{Fe}^{3+} \) compounds to oxidize ethanol into \( \text{CO}_2(g) \) and \( \text{H}_2(g) \), but in strict co-operation with methanogens [54]. Lastly, the *Myxococcales* and *Chlamydiae* species have been identified as fermenters of carbohydrates, [67] and proteins [68], which are ultimately converted into organic acids. All these δ-Proteobacterial roles contribute to the systems’ stability by providing electrons, which act as reducing equivalents in the AD systems. We also postulate that differences in the cow-dug’s substrate composition were the main drivers of the observed variations in the δ-Proteobacteria abundances and β-diversities among the treatments (SFig. 6b). The other low abundance species identified in our study were assigned to the *Mariprofundale’s* order, the affiliates of Z-Proteobacteria class. The affiliates co-exist with *Nitrosopumilales* to oxidize \( \text{Fe}^{0} \) [69] into \( \text{Fe}^{2+} \) or \( \text{Fe}^{3+} \) ions. Moreover, the two phylotype species have previously been revealed to convert nitrogen substrates into \( \text{NO}_3^{-}(aq) \) in their syntrophy [70, 71]. In this study, through their syntrophy, we hypothesize a completion of redox reactions in the treatments through electrons release. In summary, all affiliates of *Proteobacteria* identified in this study are capable of reversing oxidative damage to methionine [54], a trait that explains why their species are more (richness) compared to other phyla species in our treatments. Moreover, *Proteobacteria* species are able to express shorter diversification time-scales over other microbes hence leading to large population sizes, high growth rates, and strong selective
regimes of the species. All these factors facilitate rapid Proteobacteria adaptation through either mutation or recombination events.

We also identified Solibacterales and Acidobacterales species that were affiliated to the phyla Acidobacteria. Members of these species ferment polysaccharides and monosaccharides into acetate and propionate [72]. The species are also able to fix CO\textsubscript{2(g)} anaplerotically [73], ultimately replenishing carbon atoms in the AD systems. Furthermore, the affilaites are also able to catabolize inorganic and organic nitrogen substrates as their N-sources [74]. The detection of these species in our study is attributed to the high amount of NH\textsubscript{4(g)} in our AD systems. We postulate that reactor perturbations rather than environmental variables were the main influencers of the genetic composition of the Acidobacteria in our study (Table 3 and SFig. 25b and 26b).

Generally, the Mollicute’s species (Tenericutes phylum) detected in our study are facultative anaerobes. These were all absent in reactor 5, probably due to the nature of the substrates or feeds fed to the cattle breeds. The species are fermenters of sugars, and they convert the substrates into organic acids such as lactic acid, saturated fatty acids and acetate [75, 76]. All the identified species are affiliates of Acholeplasmatales, Mycoplasmatales and Entomoplasmatales which have previously been shown to exhibit reduced respiratory systems, with incomplete TCA cycle, that lack quinones and cytochromes. The metabolisms of Tenericutes generally yield low amount of ATP and large quantities of acidic metabolic end products [75]. It is probable that the produced acidic metabolites negatively affected the performance of other microbes in the systems in our study; which led to poor reactor performance. We hypothesize that the occurrence of these species in our systems was driven by the agroecological zones and soil types, which directly affected the nature of the cow-dug’s substrates fed into our treatments (Table 2 and 3 and SFig. 30).

Like other Bacterial species, the affiliates of Actinobacteria (SFig. 17a) identified in this study, are able to metabolize a wide variety of substrates [77, 78, 79] to produce acetic acid, lactic acid, H\textsubscript{2(g)} and other products. Among the Actinobacteria, Coriobacteriales metabolize formate and H\textsubscript{2(g)}, and concomitantly reduce NO\textsubscript{3(aq)} into NH\textsubscript{3(g)} [77]. Slackia Heliotrinireducens was one of the Coriobacteriales species identified in this study. Species of the Bifidobacteriaceae were also identified in our study. These species are able to metabolize oligosaccharides and release lactic acid [78]. Other species belonging to Actinomycetales syntrophically metabolize propionate into acetate [79]. The metabolism of NH\textsubscript{4(aq)} by Actinobacteria is known to protect methanogenic Archaea from other oxidizing agents. However, their species have been reported to be more sensitive to nutrients perturbations [79] and due to this, reactor perturbations particularly the feeding regimes (SFig. 17b) may have contributed to the observed species abundances and diversities among our treatments in this study. The low abundance of Plactomycetes in this study has also been observed elsewhere [80]. Plactomycetes have been found to metabolize sulphate-containing polysaccharides [80], recalcitrant hydrocarbons and other organic compounds [81]. However, when in syntrophic association with Chloroflexi (SFig. 18a, 19a, 20a and 21a), the species are also able to oxidize butyrate [82] into smaller substrates.

The low abundance of communities of Thermotogae in this study is similar to findings of previous studies [83]. The carbohydrate metabolism roles of the Thermotogae were previously determined based on the gene content of Defluviitoga tunisiensis [84]. The species have been reported to convert the macromolecule into acetate, ethanol, CO\textsubscript{2(g)}, and H\textsubscript{2(g)}. In most cases, the species utilize the Embden-Meyerhof Parnas (EMP) pathway to generate pyruvate, which is either converted into acetyl-CoA, a precursor for acetate and ATP
formation or reduced to ethanol, while at the same time oxidizing electron carriers. The H$_2$(g) produced in these reactions is coupled to the oxidation of NADH and reduction of ferredoxin [84]. The reactions have been found to favour proton production which reduces the $P_{H2}$ in the AD systems. However, for these reactions to proceed favourably Thermotogales have been revealed to depend on Clostridiales species [85] which implies that the abundance and richness of Thermotogale’s is regulated by the Clostridiales species. The Spirochaetes were detected in low abundance in our study as was also found out by Wang and his co-workers [86]. The identified Spirochaetes in this study are homoacetogens and acetogens which have previously been revealed to convert H$_2$(g), CO$_2$(g) and CO(g) into acetate [86, 87]. These species have also been revealed to co-exist with methanogens to provide an alternative metabolic route, termed as the butyrate oxidation pathway for H$_2$(g) consumption [87]. Similarly, the Spirochaetes species have also been found to co-exist with the Thermotogales species and to jointly ferment tetrasaccharides and cellulbiose into organic acids [88]. These syntrophies protect Clostridiale’s species from substrate inhibition mechanisms [85]. Based on our findings and those reported elsewhere, we suggest that numerous microorganisms identified in our AD systems co-operate through the use of an assortment of mechanisms to metabolize the substrates. Due to this co-operation, we further postulate the existence of a holobiontic mechanism in our treatments that is used to regulate symbionts and other syntrophic species through ecological selection, via co-evolution to minimize conflict between or among the symbiotic species.

The Verrucomicrobia communities identified in this study were affiliated to Methylacidiphiles, Punicicecoccales and Verrucomicrobiales orders (SFig. 28a and 29a). These communities have been established to metabolize cellulose and other C$_1$-containing compounds [89]. The affiliates of Cyanobacteria were classified as Chroococcales, Oscillatoriales and Nostocales species (SFig. 23a and 24a) which are mainly involved in H$_2$(g) and N$_2$(g) metabolism [90-92] in addition to carbohydrate metabolism. Research has shown that the conversion of N-containing substrates into NH$_3$(g) is an endergonic reaction that requires ATP [90-92]. The reaction has also been found to oxidize H$_2$(g) into obligate H$_2^{aq}$ ions, circumventing high $P_{H2}$ in the AD systems [90-92]. Therefore, the recycling of H$_2$(g) through the oxyhydrogen route, also termed as the Knallgas reaction by the Cyanobacteria species has been revealed to provide ATP and protect strict anaerobes and other enzymes from oxidizing agents, by providing reducing equivalents and removing toxic O$_2$(g) species [93, 94]. In our case, we attributed the variation in abundance of Cyanobacteria species among the treatments to the variability of H$_2$(g) concentration among the treatments.

Another identified phylotype in our study was the Candidatus poribacterial species. These species ferment carbohydrate and proteins, including urea [95] into organic acids. The species have been revealed to use the Wood-Ljungdahl pathway also known as CO$_2$(g) reductive pathway to fix CO$_2$(g) generated from carbohydrates and other macromolecules. The species have also been previously reported to utilize the oxidative deamination pathway [96] to ferment amino acids, and the randomized methyl-malonyl-CoA pathway to metabolize propionate. All these reactions produce H$_2$(g) and we postulate commensal symbiontism between C. poribacteria and acetogenic species in our treatments. Like other Bacteria species, Fibrobacter members such as Fibrobacter succinogenes identified in our study were likely to have been involved in the metabolism of cellulose [97] and other macromolecules in our AD systems.

Although it’s not a common tradition to explain the existence of metal-reducing organisms in the AD systems, it is worth noting that the electron sink ability of any system depends on these organisms. One of such phylotypes identified in our study were the affiliates of
**Gloeobacterales** order (SFig. 22a and 23a). The member species have previously been reported to transfer electrons directly from methanogens to syntrophic **Bacteria** [98]. These kinds of reactions are thought to enhance clean biogas production through the inhibition of sulfate reducing bacteria. Other identified metal reducing species in our study were the members of **Clostridiales** that have previously been reported to utilize Fe$^{3+}$, Co$^{3+}$, and Cr$^{6+}$ as electron acceptor [98]. Other identified **Clostridiales** species have also been reported to use the acrylate pathway to reduce arsenate and thiosulfate concomitantly fermenting organic molecules [99]. The affiliates of **Nitrospirales** (**Nitrospira** phylum) identified in this study, have been reported to reduce sulfate using H$_2(g)$ and thereby promoting acetate oxidation [100, 101]. We hypothesize that all these metal reducing reactions contributed to the reactor stability through the provision of reducing equivalents in the systems.

**The identified Archaeal biomes**

The generated volatile organic acids, CO$_2(g)$ and H$_2(g)$ by acidogens, acetogens, and homoacetogens are further metabolized by **Archaea**, particularly the species that generate CH$_4(g)$ [102]. The **Archaea** domain constituted of six (6) methanogenic and ten (10) non-methanogenic guilds identified in our study, with the two groups playing equally important roles in the AD systems. The methanogenic species are consumers of **Bacterial** products and convert them into CH$_4(g)$, the biogas. The methanogenic and non-methanogenic species identified in this study are affiliates of **Euryarchaeotes**, which utilize hydrogenotrophic, acetoclastic, methylo trophic and methyl-reducing pathways to metabolize the substrates of **Bacterial** species into CH$_4(g)$. Generally, the species of **Euryarchaeota** are known to indirectly stabilize the AD systems by consuming H$_2(g)$ [103] and we suggest that the varied H$_2(g)$ concentrations among the treatments, contributed to the observed variances in **Euryarchaeota**’s abundance and β-diversity among the local communities. However, environmental (Table 1 and 2) and reactor perturbations may also have influenced the community abundance. Among the methanogens, **Methanosarcinales**, the affiliates of **Methanomicrobia** class, were the second most abundant group, consisting of 1-3% of the total communities identified. The affiliates of **Methanosarcinales** use three metabolic routes singly or in combination to convert methanol and methylamines, acetate or CO$_2(g)$/H$_2(g)$ and other potential substrates into CH$_4(g)$. The produced H$_2(g)$ from the reactions is also consumed by the same species to maintain $P_{H2}$ [104]. However, some species within the guild, are known to lower the accumulated $P_{H2}$ through the reassimilation of acetate, a process that transfers H$_2(g)$ to acetogens. Under certain conditions, the species of **Methanosarcinales** (e.g **Methanosarcina acetivorans** C2A and **Methanosarcina barkeri**) and **Methanobacteriales** (**Methanothermohacterthermoautotrophicus**) have been reported to express metabolic features similar to aceticogenic organisms. Previously, these species have been reported to use the acetyl-CoA pathway to metabolize CO$_2(g)$ into acetate and formate, under different $P_{CO_2(g)}$ [105]. These reactions provide an alternative route for ATP generation from CO$_2(g)$ reduction by the generated H$_2(g)$. Normally, the **Methanosarcinales** couple methanogenesis to ion transport, also termed as chemiosmotic coupling or chemiosmosis, to establish an electrochemical gradient across the cell membrane [105] because the CO$_2(g)$ reduction reaction is normally unfavourable. The **Methanosarcinales** have however been reported to express significant metabolic plasticity.

Among the **Archaea** in our study, **Methanomicrobiales**’ affiliates of the **Methanomicrobia** class (SFIG. 36a) were the most abundant within the class, comprising of 2-4% of the identified total microbial populations. The affiliates of **Methanomicrobiales** use two methanogenesis pathways; hydrogenotrophic and methylo trophic routes, to convert formate, H$_2(g)/$CO$_2(g)$, 2-propanol/CO$_2(g)$ and secondary alcohols [106] into CH$_4(g)$. The **Methanomicrobiales** consume isopropanol to yield acetone and the reducing equivalent,
NADPH [107]. However, the species are known to lack the CO$_2$(g) reductive pathway, and hence uncouple methanogenesis from CO$_2$(g) fixation, which implies that the species uses electron birfucation to conserve energy. Due to these characteristics, the *Methanomicrobiales* species are able to express the high acetate requirement phenotype for carbon assimilation [108]. In contrast, the identified *Methanocellale*’s species, which are affiliates of the same class, solely use the acetoclastic route to convert CO$_2$(g) into CH$_4$(g) [109]. Majority of the *Methanocellale*’s species have been reported to reduce CO$_2$(g) with H$_2$(g), except in rare cases where formate and secondary alcohols are used as alternative electron donors. Similarly, all the identified *Methanobacteriales*’ species, the affiliates of *Methanobacteria* class use the CO$_2$(g) reductive route to convert CO$_2$(g) into CH$_4$(g) with formate and H$_2$(g) acting as electron donors. We are prompted to postulate that formate metabolism particularly in acetoclastic methanogenesis was triggered by the low $P_{H2}$ and that through electron bifurcation the species were able to conserve their energy. Other detected methanogenic species were the affiliates of *Methanopyrales*, belonging to *Methanopyri* class and *Methanococcales* species that were assigned to *Methanococci* class (SFig. 34a and 35a). The members of the two classes have previously been reported to metabolize H$_2$(g) and CO$_2$(g) into CH$_4$(g) under thermophilic and saline conditions, [110, 111], which may indicate that our treatments had high amounts of salts and that the temperatures were high enough to support the species.

In contrast, the non-methanogenic group of *Archaea* generally reduce metal elements and their compounds [98, 99], to provide reducing equivalents in the AD systems. The affiliates of *Halobacteriales* (*Halobacteria, Euryarchaeota*) were among the detected non-methanogenic species in this study. The member species of this group of non-methanogens use the arginine deiminase pathway [112, 113] or other metabolic routes to catabolize amino acid substrates. The *Halobacteriales* conserve energy when the intermediate metabolites; ornithine and carbamoyl-phosphate, are further converted into CO$_2$(g) and NH$_3$(g). The significant differences in relative abundances of the *Halobacteriales* among the treatments can be attributed to the concentration of the alternative electron acceptors and salts concentrations in the systems. However, other deterministic factors such as competition and niche differentiation could also have led to the observed variations in *Halobacteriales* abundance among the treatments.

Our study also identified *Thermoplasmatales*, the affiliates of *Thermoplasmata* class. These groups of *Archaea* utilize sulfur for energy conservation [114]. However, their close relatives, *Methanomassilicoccales*, formerly *Methanoplasmatales* have been found to use the methyl-reducing pathway to convert methanol into CH$_4$(g). These relatives conserve energy through electron bifurcation [115] and we postulate that the affiliates of *Thermoplasmatales* are also involved in CH$_4$(g) production although few or none of the biochemical assays have previously been conducted to test their biogas production abilities. We also identified *Thermococcales* species, of the class *Thermococci*. The affiliates of this class of *Archaea* have been reported to have short generation time and to withstand nutrient stress for longer durations [116]. Like for *Bacterial* species, and in the absence of sulfur, *Thermococcales*, have been reported to be involved in both acidogenesis and acetogenesis reactions. Species of this group metabolize proteins and carbohydrates into organic acids, CO$_2$(g), and H$_2$(g) [117, 118]. Other affiliates of *Thermococcales*’ are involved in homoacetogenic reactions through which they oxidize CO$_2$(g) into CO$_2$(g), and catabolize pyruvate into H$_2$(g) and other products [119, 120]. We also identified *Archaeoglobales*, the affiliates of *Archaeoglobus* class which use sulfate, as the reducing agent, to oxidize lactate, long fatty acids and acetate into CO$_2$(g) [121]. Some species e.g *Geoglobus ahangari*, however, reduce Fe$^{3+}$ compounds with H$_2$(g) [122]. The identified *Archaeoglobales* and other metal reducing species identified in our
study indicate the importance of these microbes in redox balancing reactions which contribute to reactor stability.

The *Sulfolobales* and *Desulfurococcales*, all belonging to *Thermoprotei* and *Crenarchaeota* phylotypes were also identified in this study (SFig. 37a). Species of these groups metabolize carbohydrates and amino acids [123, 124] into pyruvate and CO$_2$. They also ferment pyruvate into lactate and CO$_2$. The CO$_2$ generated from the above reactions either enters into the CO$_2$ reductive pathway, the route that is catalyzed by acetogens/homoacetogens to produce CH$_4$ or is fixed through the 3-hydroxypropionate/4-hydroxybutyrate or dicarboxylate/4-hydroxybutyrate cycles by the identified *Thermoprotei* species to generate the energy [125]. Among other *Thermoprotei* affiliates identified in this study were *Thermoproteales* that use sulfur and its compounds as electron acceptors, to ferment carbohydrates, organic acids and alcohols [126]. In these reactions, the species consume CO$_2$ and H$_2$. Though the species have been previously reported to grow at low rate, they also ferment and oxidize amino acids and propionate, respectively [127]. Other affiliates like *T. neutrophilus* are known to assimilate acetate and succinate in the presence of CO$_2$ and H$_2$ [128]. This species use H$_2$ to reduce sulfur. Unlike most of the other *Archaeal* species, *Thermoproteales* use oxidative and reductive TCA routes to conserve energy. *Acidilobales* are another *Thermoprotei* order identified in this study uses S$^0$ and thiosulfate to metabolize protein and carbohydrate substrates [129]. The species use either EMP or the Entner-Doudoroff (E-D) [130] pathways to convert the substrates into acetate and CO$_2$ [131]. The affiliates of *Acidilobales* are also able to oxidize triacylglycerides and long-chain fatty acids [131] into acetate and CO$_2$. Substrate-level and oxidative phosphorylation are the major routes of energy conservation in *Acidilobales*; and the two mechanisms occur when S$^0$ and acetyl-CoA concentrations are high. Due to S$^0$ respiration in *Acidilobales* species, we postulate that their member species close the anaerobic carbon cycle via complete mineralization of organic substrates and the variations of these substrates together with other deterministic and non-deterministic (Table 2 and 3) factors contributed to the observed variation in abundance among the treatments and the high β-diversity (SFig. 37b). We also hypothesize that soil type was another factor that exerted natural selection on the *Acidilobales* through the alteration of the cow-dung’s substrates chemical composition.

The identified *Thaumarchaeota* communities were classified into *Cenarchaeles* [132] and *Nitrosozoomilales* [133] phylotypes (SFig. 38). The member species of these groups use the modified 3-hydroxypropionate/4-hydroxybutyrate pathway to fix inorganic carbon to generate energy. The communities are also able to metabolize urea and its products into, NH$_3$ [134, 135]. The *Nanoarchaeota* (SFig. 31a) species exploit *Crenarchaeota* mechanisms [136, 137] to metabolize amino acids and complex sugars [138, 139] while the *Korarchaeotes* are sole amino acid fermenters. The energy metabolism in the last two phyla is still unclear and further biochemical studies are needed to unravel their full metabolic activities.

**The identified fungal biomes**

Like the *Bacterial* species described in this study, the identified *fungi* (SFig. 39a, 40a and 41a) have the ability to convert macromolecular substrates into acetate, CO$_2$ and H$_2$ [140]. Among the identified fungal species, *Ascomycota* were the richest and most abundant guild, comprising of <1% of the identified communities in the treatments. Among the *Ascomycota* species, a majority were affiliates of *Saccharomycetales* (*Saccharomycetes, Ascomycota*), that generally use the EMP pathway to ferment glucose and other sugars into pyruvate. One of such identified species is *Saccharomyces cerevisiae* [141] that is known to ferment carbohydrates into mixed fermentation products. Other *Saccharomycetes* species e.g
Hansenula polymorpha yeast ferment monosaccharides, fatty acids and amino acids into acetate, ethanol, CO$_2$(g) and H$_2$(g) [142]. Furthermore, the species are also able to oxidize methanol into formaldehyde and H$_2$O$_2$(g) [143]. The two products act as the sources for formate and H$_2$(g), the precursors for methanogenesis. The observed variations in abundance of the Saccharomycetes species may be attributed to their metabolic plasticity, deterministic factors and other environmental variables. The other Ascomycota affiliates belonging to Pezizales (Pezizomycetes), Pleosporales (Dothideomycetes), and Eurotiales (Eurotiomycetes) identified in this study use mixed fermentation pathways to ferment carbohydrates [144, 145]. The metabolisms of these Ascomycota species have however been observed to vary within and among the phylotypes, despite the presence of a common enzyme for the pyruvate metabolism [144, 145]. For example, among the Eurotiale’s order; the species Paecilomyces variotii expresses a low metabolic capacity on plant-derived compounds compared to other Eurotiales’ members [146, 147]. This seems to indicate the genetic diversity of Eurotiomycetes’ species, seen in the variations of abundance of the species among the treatments in this study. Our data also suggests that the communities of Eurotiomycetes were influenced by the environmental temperature and soil type (Table 2 and 3 and SFig. 45a). Like the identified Fusobacteria species and other identified amino acid metabolizers, Dothideomycetes, another amino acid consumer phylotype, metabolize lipids and proteins into volatile organic acids [147]. On the other hand, the affiliates of Pezizomycetes use sulfate as an electron sink to consume amino acids [148, 149]. We attribute the community variation observed within the above groups to the composition of the cow-dung’s substrates and natural selection pressure differences among the treatments.

The Sordariomycetes’ affiliates identified in this study (SFig. 44a) produce enzymes that express an assortment of degradative mechanisms and other fermentative roles. The Hypocreales’ affiliates are such an example that produces powerful hydrolytic enzymes [150] that catabolize complex biomass in their vicinity. Fusarium sp. a member of Hypocreales, metabolizes NH$_3$(g), lipids and proteins [151, 152] into the respective smaller molecules. The species has been found to use NAD(P)H to reduce NO$_3$$_{(aq)}$ into NH$_4$$_{(aq)}$. The production of NH$_4$$_{(aq)}$ by the Hypocreales is normally coupled to the substrate-level phosphorylation and ethanol oxidation; the processes that produce acetate [151, 152]. The two mechanisms ameliorate the NH$_4$$_{(aq)}$ toxicity to the micro-organisms. Other Hypocreales species such as Hirsutella thompsonii confer wide metabolic activities that involve chinases, lipases, proteases and carbohydrate degrading enzymes [153]. Monilia sp. [154] and Neurospora crassa [155] were among the identified Sordariomycetes and they convert cellulose into ethanol. Magnaporthales’ such as Pyricularia pennisetbi convert isovaleraldehyde into isovaleraldehyde while other Magnaporthales’ species convert fumaryl-acetoacetate into fumarate and acetoacetate [156, 157]. Other roles played by the Sordariomycete’s species include the metabolism of aromatic compounds through which, phenolics and aryl alcohols/alddehydes are converted into acetate, formate and ethanol [158, 159]. Our data suggests that the Sordariomycete’s community’s abundances were influenced by the cow-dug’s substrates and the environmental variables (Table 2 and 3) of the reactors.

In contrast to the Ascomycota species, affiliates of Basidiomycota establish susceptible hyphal networks [160] that may have been constrained by the reactor pertubartions in our study. This may explain why the Ascomycota species dominated over the Basidiomycota’s species in this study. Nonetheless, like Ascomycota species, the affiliates of Malasseziales, Ustilaginales, Agaricales and Tremellales species, all belonging to Basidiomycota, are able to metabolize carbohydrates [145, 161], urea [161] and NO$_3$$_{(aq)}$ [161] in the systems. Other members like Polyporales metabolize woody substrates and their respective products [162]. All these functions clearly indicate that Basidiomycota’s species played diverse metabolic
roles in our AD systems in this study. Interestingly, we also identified Chytridiomycota affiliates in reactor 9 and 10. These affiliates are capable of catabolizing pollen, chitin, keratin, and cellulose [163, 164], which may be indicative of the presence of these substrates in the two treatments. The detected Spizellomycetales, Cladochytriales and Rhizophydiales species in the two treatments, further revealed the presence of highly decayed plant tissue with high cellulose content. The identified Monoblepharidomycetes members are close relatives of Cladochytriales and Spizellomycetales [165] species and were detected only in reactor 9 and we postulate that these groups perform similar roles in the AD systems.

Generally, Chytridiomycota species are carbohydrate metabolizers as previously revealed by the presence of glucose transporters and invertase activity in their systems [163, 164]. The species use electron bifurcation mechanisms to conserve their energy and the mechanisms occur in the hydrogenosome [166]. The Blastocladiales, the close relatives of Chytridiomycotes and the affiliates of Blastocladiomycota, class Blastocladiomycetes were only detected in reactor 9 in our study. Their member species metabolize cellulose into acetate, CO\(_2\)(g) and H\(_2\)(g). Interestingly, Blastocladiales have previously been reported to use the Malate fermentation pathway to generate and conserve their energy [167]. Mortierellales formerly classified under Mucoromycotina [168] and latter assigned to Mortierellomycotina phylum [169] are converters of organic compounds into poly-unsaturated fatty acids [170] while their relatives Entomophthorales and Mucorales metabolize decaying organic matter. However, the affiliates of the three orders were only detected in reactor 9, 10 and 11, respectively.

Like similar studies [171, 172], our data revealed that fungi co-occurred with the species of Bacteria and methanogens. In these syntrophy, we postulate that the fungal species catalyzed inter-species H\(_2\)(g) transfer and re-generated oxidized reducing equivalents: NAD\(^+\) and NADP\(^+\) in the systems. All these reactions increased metabolism of the dry matter [173] in the treatments. We also suggest that the fungal syntrophy is more complex than simple cross-feeding mechanisms, because the species shift more oxidized products into a more reduced form. Through these mechanisms, the precursors, lactate and ethanol are converted into acetate and formate depending on the prevailing conditions. The involved interactions between fungi and other domain’s species have been determined to be very crucial to an extent of influencing the isolation and culturing of some fungal species [173, 174]. The findings indicate why fungal species should be jointly considered with other domains’ species in the context of biogas production.

**The microbiomes’ α- and β-diversities and biogas production**

To establish whether environmental conditions and perturbations influenced the microbial populations, we compared the collected phenotypic traits with the genotypic dataset. The phenotypic datasets revealed an insight on the cause of the observed communities’ β-diversities and biogas production variation. We attempted to use the principles of the four ecological mechanisms that include selection, drift, speciation, and dispersal to better understand the cause of the observed variance in microbial composition among the treatments and thus the cause of the biogas production variability among the sites. While selection is deterministic and ecological drift is stochastic, dispersal and speciation may have contributed to the two ecological parameters. One of the challenges of examining communities’ in any given niche is the difficulty of estimating real dispersal among the sites. Nonetheless, our treatments were fluidic, naturally encouraging rapid microbial dispersion. Hence, species dispersal was not the major limiting factor that influenced our local communities. Similarly, we also excluded speciation based on our sampling process and experimental design. Therefore, we suggest that selection and drift were the likely two main factors that
contributed to variation of community abundances among the treatments and the observed high β-diversities in our study. The two factors were most likely influenced by the substrate inputs (Table 3) and disturbances from environmental variables (Table 2 and 3). Environmental conditions directly affect the chemical composition of the substrate fed to the cattle breeds. As a result, their wastes, the cowdung is fed to the reactor which ultimately influences the microbial composition and biogas productivity. Temperature, agroecological zones and soil types were among the environmental factors that were observed to drive the specific communities’ β-diversities and their relative abundances among the treatments. All these could have contributed to the observed variation of biogas production among the studied treatments. We postulate that environmental temperatures combined with other deterministic factors such as the species metabolic traits, largely contributed to the lack of identification of the Chytridiomycota, Blastocladiomycota and other rare fungal species in the majority of the treatments. The findings increased the number of the total species in the few treatments, where the rare species were identified as opposed to where they were absent. Nonetheless, we never observed significant differences in reactor performances between the two categories of the treatments. Consequently, we linked the observations with the low abundances of the rare fungal species identified in the few respective reactors, which implies the importance of the species abundance in biogas production. The composition of the cowdung’s substrates fed to these treatments could largely be influenced by cattle breeds and several environmental factors including agroecological zones, soil type, rainfall, environmental temperatures to mention but a few, and its chemical compositions could have stimulated the growth of specific species, favouring colonization and speciation in the specific treatments. Similarly, the same phenomenon could have happened to other microbial communities. In this case, the available resources could have reduced competition among the species, weakening the niche selection abilities. Generally these ecological factors normally strengthen the ecological drift and priority effects [175] that result to unpredictable site to site variation of microbial abundances and high β-diversities, even under similar environmental conditions. We used the same arguments to explain our observations on the variations of the relative abundances among treatments and high β-diversities, particularly when the conditions were similar (Table 2 and 3). We also extended the same argument to the biogas production capacities among the treatments.

Besides the cow-dung’s substrate effects, we hypothesize that the intensity and duration of the disturbances [176], led to high rate of species mortality, low growth rates and species extinction. These kinds of disturbances caused decrease of stochasticity and to account for their effects on our findings, we adopted the concepts of neutral theory [176]. Generally their associated caveats reduce ecological drifts that result to low competition abilities among the low abundance species. Due to these facts, we postulate that the persistence and intensities of these disturbances in our specific treatments eliminated weak and less adopted microbial populations. One of such identified case was revealed in reactor 5 in which the Tenericutes affiliates were absolutely extinct. Similarly the absence of Mortierellales, Entomophthorales, Mucorales to mention but a few, in other treatments (SFIG. 51) was an indication of unfavorable disturbances that could have originated from either cow-dung’s substrates or other reactor perturbations. All these may have led to more predictable site-to-site variations and we used these concepts to explain the observed low β-diversity among the treatments (e.g in reactor 2 and 10, reactor 3 and 6 etc). However, to understand better the links between the environmental conditions and the microbial communities in these AD systems, we suggest the utilization of the metacommunity theory that considers the world as a collection of patches that are connected to form a metacommunity through organisms’ dispersal. The model depends on the specific species traits, the variance of the patches in the environmental
conditions as well as frequency and extent of species dispersal in the systems. Other ecological models lack these features. Therefore, the model can be used to ask and test whether dispersal, diversification, environmental selection and ecological drift were the vital elements in the biogas production process. In this case, a detailed understanding of the contribution of these factors on specific species abundances and community assemblage can be used to tailor the reactor management from an ecological point of view.

**Conclusions and Future prospects**

Our results provide an insight into the biogas microbiome composition, richness, abundances and variations within and among the taxa and the twelve reactor treatments. Interestingly, microbiome structure and richness were similar among the twelve treatments, except for the affiliates of *Rhizophydiales*, *Cladochytriiales*, *Monoblepharidales*, *Blastocladiales* and *Entomophthorales*, which were detected only in reactor 9. Contrasting the findings were the affiliates of *Mortierellales* and *Mucorales* that were identified in reactor 10 and reactor 11, respectively, while the *Spizellomycetales* species were detected in both reactor 9 and 10. We also revealed that the affiliates of *Tenericutes*, were the only *Bacterial* guild that was absent in reactor 5. The Cow-dung’s substrates and metabolic capabilities of these species were the main elements that were considered to limit their dispersal to other treatments. Nonetheless, the identified phylotypes richness within and among the phyla were found to vary, except for the phyla, that comprised a single phylotype. All the identified species abundances among the treatments and the community’s β-diversities in the majority of the treatments significantly varied. All these contributed to our conclusion that environmental and spatial variables largely influenced biogas microbiome populations and thus the amount of biogas production. The discoveries highlight the importance of identifying the involved natural processes that give rise to such identified variations. Based on the results obtained in this study, we suggest further studies on species responses on the natural conditions for optimal biogas production and microbial community managements. We also suggest the application of metacommunity theory among other ecological theories in developing management strategies for the AD systems.

**Materials and Methods**

**Ethics and experimental design**

The sampled treatment sites (reactors) were installed in the premises of privately owned land (Table 2) and permission to sample was granted by the owners. Randomized complete block design was utilized as previously described [177] to assign the twelve cowdung digesting treatments to the four blocks, which operated at mesophilic conditions. Organic loading rate and the hydraulic retention time were largely influenced by the the treatment size and the farmers’ routine activities. Sampling was conducted following the previously described procedure of Boaro et al. [178]. Geographic co-ordinates (Table 2) and agroecological zones (Table 3) were generated by Geographic Positioning System (etrex, SummitHC, Garmin) and ArcGIS® software (ver10.1) [179], respectively. A pre-tested questionnaire was utilized according to Oksenberg et al. [180] protocol to collect the phenotypic dataset. Small scale farming and zero grazing of Freshian breeds were the main activities among the sampled farmers. However, the owners of reactor 3 and 11 (Meru and Kiambu blocks, respectively), reared Jersey breeds of cattle. Samplings were conducted during the warm and cool climatic conditions.

**Sample preparation and Sequencing**

The double stranded deoxyribonucleic acids (dsDNA) was extracted from the duplicate sludge samples using ZR soil microbe DNA kit (Zymo Research, CA., USA), according to manufacturers instructions. The quantity and quality of the extracted dsDNA were determined as previously described by Campanaro et al. [181]. The dsDNA was normalized
and library preparation performed using Nextera™ DNA sample prep kit (Illumina, San Diego, CA, USA), according to manufacturer's instructions. Library normalization was conducted using Illumina TruSeq DNA PCR-free sample preparation kit buffer (Illumina, USA). To prepare for cluster generation and sequencing, equal volumes of normalized libraries were pooled, diluted in hybridization buffer and denatured, according to Nextera™ protocol. De novo deep sequencing was conducted by an Illumina MiSeq system as previously described by Caporaso et al. [182] but adopting a v3 kit chemistry with 300 cycles (2×150) as previously described [181].

**Bioinformatics and Statistical analysis**

All paired-end reads in FASTQ format were pre-processed with FastQC software (ver0.11.5) following the procedure described by Zhou et al. [183]. Trimmomatic software was utilized to trim and filter the adaptor according to [184]. De novo assembly of the paired-end contigs were performed by MetaSpade software (ver3.10) as described previously [185]. To re-trim low-quality regions of the scaffolds, we utilized SolexaQA as described by Cox et al. [186] while DRISEE, metagenomic quality assessment software was used to remove duplicated scaffolds according to Keegan et al. [187]. We screened near-exact scaffolds that matched human, mouse and cow genomes following Langmead et al. procedure that used a sort read aligner, the Bowtie [188], M5nr [189], a non-redundant tool, FragGeneScan [190] and sBLAT [191] tools were utilized to compute datasets against 16 reference biological databases as previously described by Wilke et al. [189]. Two ENVO ontologies: large lake and terrestrial biomes were utilized according to Field et al. [192] and Glass et al. [193]. Alignment similarity cutoff were set at single read abundance; ≥60% identity, 15 amino acid residues and maximum e-value of 1e-5. The annotated datasets were archived in CloVR [194]. Prior to statistical analysis, the rarefaction curves were reconstructed, datasets was log2 (x+1) transformed and normalized using an R package, DESeq software [195]. PCoA based on euclidean model [196] was set at \( P \leq 0.05 \), and conducted against the twelve treatment samples. Results were visualized by MG-RAST pipeline (ver4.01) [197] and Krona radial space filling display [198]. Optimal scripts and restful application programming interface were utilized in our analysis.
### Table 2: The Geographic positioning systems’ co-ordinate for the sampled reactor treatments, assigned to the four blocks of eastern and central regions of Kenya

| Reactor Number | Sample id | Blocks | Treatment | Northings | Eastings | Altitude | Location |
|----------------|-----------|--------|-----------|-----------|----------|----------|----------|
| EA12/11/09/004 | S_1       | Kiambu | L-C       | S(00'15.477") | E(37.38.828") | 1468 M   | Ngenda   |
| EA12/10/07/003 | S_2       | Kiambu | S-C       | S(00.21347") | E(37.61867") | 1674 M   | Karimoni  |
| EA12/10/04/001 | S_3       | Meru   | S-E       | S(00'08.180") | E(37.39.477") | 1443 M   | Kaithe    |
| CE20/12/01/020 | S_4       | Kirinyaga | S-C     | S(00'32'45.97") | E(37'26'40.52") | 1364 M   | Murinduko |
| EA12/10/08/004 | S_5       | Meru   | L-E       | N(00'01.997") | E(37.35.353") | 1958 M   | Kithaku   |
| EA12/10/09/002 | S_6       | Meru   | M-E       | N(00'04.046") | E(37.38458") | 1650 M   | Ntima     |
| EA12/11/02/002 | S_7       | Meru   | L-E       | N(00'29.046") | E(37.10077") | 1987 M   | Ontulili   |
| EA14/11/05/001 | S_8       | Embu   | L-E       | S(00'03.269") | E(37.28.247") | 1306 M   | Municipality |
| EA14/11/09/001 | S_9       | Embu   | M-E       | S(0'27.824") | E(37.32.537") | 1417 M   | Gatari South |
| CE12/10/06/012 | S_10      | Kiambu | L-C       | S(00'24.874") | E(37.62438") | 1589 M   | Kianyu    |
| CE12/10/07/013 | S_11      | Kiambu | S-C       | S(00'16'17.57") | E(37.38'40.16") | 1577 M   | Dururumo  |
| CE14/11/11/001 | S_12      | Kirinyaga | M-C     | S(00.48637") | E(37.44057") | 1539 M   | Ngandori  |

**Legend:** CE: Central regions, EA: Eastern regions, L-C: Large size, Central regions, M-C: Medium size, Central region, S-C: Small size, Central regions, L-E: Large size, Eastern regions, M-E: Medium size, Eastern regions S-E: Small size, Eastern regions.

### Table 3: The ArcGIS Software and the Questionnaire generated environmental variables and Phenotypic traits for the sampled treatments within the identified blocks of Kenya

| Sample id | Alt | S-T | AEZ | Ann. Av. R-F | Ann. Temp °C | Mean | Substrate source | Family size | Reactor Performance |
|-----------|-----|-----|-----|-------------|-------------|------|-----------------|-------------|---------------------|
| S_1       | 1468 M | Lithosol | UM 4 | 800-900 | 20.7-19.9 | 20.7-19.9 | CO PO PI M F T | IG ACG RV FACoG GGV |
| S_2       | 1674 M | Ferralsol | UM 2 | 1100-1400 | 19.5-18.4 | 19.5-18.4 | 3 - - 2 1 3 | 2 4 2 0 8 16 M3 14 M3 2 |
| S_3       | 1443 M | Nitosol | UM 3 | 1400-2200 | 20.6-19.2 | 20.6-19.2 | 11 - - 3 4 7 | 2 4 2 0 8 16 M3 5 M3 3 |
| S_4       | 1364 M | Andosol | UM 2 | 1220-1500 | 20.1-19.0 | 20.1-19.0 | 6 - - 4 2 6 | 1 8 1 5 6 M3 4 5 M3 1.5 |
| S_5       | 1958 M | Phaeozem | L H 1 | 1700-2600 | 17.4-14.9 | 17.4-14.9 | 3 - - 3 1 4 | 4 8 4 0 16 M3 12 M3 4 |
| S_6       | 1650 M | Nitosol | UM1 | 1650-2400 | 19.2-17.6 | 19.2-17.6 | 2 30 - 2 2 4 | 3 0 2.5 10 M3 6 M3 4 |
| S_7       | 1987 M | Phaeozem | LH 1 | 1700-2600 | 17.4-14.9 | 17.4-14.9 | 5 - - 5 2 7 | 4 8 4 0 16 M3 10 M3 10 M3 6 |
| S_8       | 1306 M | Nitosol | UM 3 | 1000-1250 | 20.7-19.6 | 20.7-19.6 | 5 - - 6 5 4 9 | 4 8 4 0 16 M3 11 M3 5 |
| S_9       | 1417 M | Andosol | UM 2 | 1200-1500 | 20.1-18.9 | 20.1-18.9 | 2 - - 3 2 5 | 3 0 2.5 10 M3 8 M3 2 |
| S_10      | 1589 M | Acbrisol | UM 2 | 1100-1400 | 19.5-18.4 | 19.5-18.4 | 3 - - 2 2 4 | 4 8 4 0 16 M3 13 M3 3 |
| S_11      | 1577 M | Cambisol | UM 3 | 800-1200 | 19.9-19.5 | 19.9-19.5 | 14 - - 5 4 9 | 2 4 2.0 8 M3 6 M3 2 |
| S_12      | 1539 M | Ferralsol | UM 1 | 1400-1700 | 19.3-17.5 | 19.3-17.5 | 3 - - 4 3 7 | 3 0 2.5 10 M3 8 M3 2 |

**Key:** Alt: Altitude (M amsl); S-T: Soil type; AEZ: Agro-ecological zones; UM 3: Marginal Coffee zone; UM 1: Coffee-Tea Zone; UM 2: Main coffee zone; LH 1: Tea-Dairy Zone; R-F: Rainfall (mm); Temp: Temperature (°C); CO: Cow; PO: Poultry; PI: Pigs; M: Male; F: Female; T: Total; IG: Ideal gas (M3); ACG: Actual calculated gas (M3); RV: Reactor Volume (M3); FACoG: Farmers Actual Collected gas (M3); GGV: Difference between the RV and FACoG.
**Acknowledgement:** We appreciate Vincent Njunge, Collin Mutai, Tina Kyalo, Joyce NJoki Njuguna, Eunice Machuka, Muga Fredrick Nganga and Philis Emilda Ochieng of BecA-ILRI hub and Donald Otieno and Linnet Gohole of University of Eldoret for their technical support.

**Funding:** This work was supported in parts by the BecA-ILRI-hub through the Africa Biosciences Challenge Funds (ABCF) fellowship Program and University of Eldoret through the Annual research grants. The funders of ABCF fellowship had no role in experimental design, data collection and processing, manuscript preparation and publishing.

**Competing Interests:** The authors have declared that no competing interests exist.

**Authors Contribution**

Conceived idea: MSM

Sourced funds: MSM

Designed experiments: MSM, OSO

Performed experiments: MSM

Performed Sequencing: MSM, MLW

Analyzed the data: MSM

Drafted manuscript: MSM

Reviewed the manuscript: MSM, OSO, PR, WFN, OJM, NPM

**Supplementary Material**

STable 1: The quality control statistics of the obtained scaffolds for the twelve treatments

SFig. 1: Bar charts showing unfiltered and filtered sequencing reads (a) and the known and unknown protein genes (b) in our samples.

SFig. 2: Stacked barchart showing some (21 phyla) Bacteria domain phyla, relative abundances (a) and the PCoA plots (phylum level), based on the Euclidean model (b).

SFig. 3: The PCoA plot showing the relative abundances variation at the class (a) and order level (b).

SFig. 4: The stacked barchart revealing six classes of *Proteobacteria* communities, relative abundances (a) and their PCoA plots based on the Euclidean model (b).

SFig. 5: The PCoA plot based on the Euclidean model for all the identified *Proteobacteria* orders.

SFig. 6: Stacked barchart showing seven *δ-Proteobacteria* orders, relative abundances (a) and their PCoA plots revealing nucleotide composition variations, based on the Euclidean model (b).

SFig. 7: Stacked barchart showing fourteen *γ-Proteobacteria* orders (a) and their PCoA plot based on the Eucleaden model (b).

SFig. 8: Stacked barchart showing seven *β-proteobacteria* orders, relative abundances (a) and their PCoA plot revealing nucleotide composition (dis)similarities among the reactors (b).

SFig. 9: Stacked barchart showing the seven *α-Proteobacteria* orders, relative abundances (a) and their PCoA plots, revealing the nucleotide composition variation among the reactors based on the Euclidean model (b).

SFig. 10: Stacked barchart showing two *ε-Proteobacteria* orders, relative abundances (a) and their PCoA plot based on Euclidean model (b).

SFig. 11: The stacked barchart revealing four *Firmicutes* classes, the relative abundances (a) and their PCoA plots, revealing variance among the twelve reactors based on the Euclidean model (b).

SFig. 12: Stacked barchart showing eight *Firmicute’s* orders, relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 13: Stacked barchart showing the four *Clostridia* orders (a) and their PCoA plots, revealing nucleotide composition variations among the twelve reactors based on the Euclidean model (b).

SFig. 14: Stacked barchart showing two *Bacillli* orders, relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 15: The stacked barchart revealing four *Bacteroidetes* classes, relative abundances (a) and their PCoA plots revealing nucleotide composition variation, based on the Euclidean model (b).

SFig. 16: Stacked barchart showing the four *Bacteroidete’s* orders, the relative abundances (a) and
their PCoA plot based on the Euclidean model (b).

SFig. 17: Stacked barchart showing six *Actinobacteria* orders, the relative abundance (a) and their PCoA plot based on the Euclidean model (b).

SFig. 18: The stacked barchart showing four *Chloroflexi* classes, relative abundances (a) and PCoA plot revealing variance among the twelve reactors, based on the Euclidean model (b).

SFig. 19: Stacked barchart showing five *Chloroflexi* orders, relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 20: Stacked barchart showing two *Chloroflexi* class orders, relative abundances (a) and their PCoA plot based on their Euclidean model (b).

SFig. 21: Stacked barchart showing the two *Thermomicrobia* orders, relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 22: Stacked barchart showing *Cyanobacteria* class, relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 23: Stacked barchart showing the five *Cyanobacteria* orders, relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 24: Stacked barchart showing the four affiliates of the Unclassified *Cyanobacteria* nucleotide reads, relative abundances (a) and their PCoA plots based on the Euclidean model (b).

SFig. 25: The stacked barchart showing the two *Acidobacteria* classes, relative abundances (a) and their PCoA plots for their nucleotide composition based on the Euclidean model (b).

SFig. 26: The stacked barchart showing the two *Acidobacteria* orders, relative abundances (a) and their PCoA plots for their nucleotide composition based on Euclidean model (b).

SFig. 27: Stacked barchart showing two *Deinococcus-Thermus* orders, the relative abundances (a) and the PCoA plots based on the Euclidean model (b).

SFig. 28: Stacked barchart showing *Verrucomicrobia* classes, relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 29: Stacked barchart showing three *Verrucomicrobia* orders, relative abundances (a) and their PCoA plot based Euclidean model (b).

SFig. 30: Stacked barchart showing three *Terneicutes* orders, relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 31: The stacked barchart showing the four *Archaea* phyla, relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 32: Stacked barchart showing nine *Archaea* classes, and the affiliates of the unclassified reads, relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 33: Stacked barchart showing sixteen *Archaea* orders, relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 34: The stacked barchart showing the eight *Euryarchaeota* classes, the relative abundances (a) and their PCoA plots revealing dissimilarities of the nucleotide composition based on the Euclidean model (b).

SFig. 35: Stacked barchart showing ten *Euryarchaeota* orders, the relative abundances (a) and their PCoA plot based Euclidean model (b).

SFig. 36: The stacked barchart showing three *Methanomicrobia* orders, the relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 37: Stacked barchart showing four *Thermoprotei* orders, relative abundances and their PCoA plot based on the Euclidean model (b).

SFig. 38: Stacked barchart showing two *Thaumarchaeota* order, the relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 39: Stacked barchart showing five fungal phyla, the relative abundances (a) and their PCoA plots based on Euclidean model (b).

SFig. 40: Stacked barchart showing thirteen fungal classes, relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 41: Stacked barchart showing 23 fungal orders, relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 42: The stacked barchart showing the five *Ascomycota* classes, relative abundances (a) and their PCoA plots based on the Euclidean model (b).

SFig. 43: The stacked barchart revealing eleven *Ascomycota* orders, relative abundances (a) and their
PCoA plot based on the Euclidean model (b).

SFig. 44: Stacked barchart showing three Sordariomycetes orders, relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 45: Stacked barchart showing the two Eurotiomycetes orders, the relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 46: Stacked barchart showing four Basidiomycota classes, the relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 47: Stacked barchart showing four Basidiomycota orders, relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 48: Stacked barchart showing the proportion of the Agaricomycetes order, relative abundances (a) and their PCoA plot based on the Euclidean model (b).

SFig. 49: Stacked barchart showing the two Chytridiomycota classes, and their relative abundances in the two reactors.

SFig. 50: Stacked barchart showing the four Chytridiomycota orders, and their relative abundances in the two reactors.

SFig. 51: Stacked barchart showing the three orders affiliated to unclassified fungal nucleotide reads.

SFig. 52: The PCoA analysis revealing β-diversity of the twelve reactors at the three taxa level: a) Domain, b) Phylum, and c) Class level.

References

1. Calusinska M., Goux X., Fossépré M., Muller E., Wilmes P. and Delfosse P. (2018). A year of monitoring 20 mesophilic full-scale bioreactors reveals the existence of stable but different core microbiomes in bio-waste and wastewater anaerobic digestion systems. *BMC Biotechnol.*, **11**:196.

2. Weiland P. (2010). Biogas production: Current state and perspectives. *Appl. Microbiol. Biotechnol.*, **85**: 849-860.

3. Verstraete W. (2015). The manufacturing microbe. *Microb. Biotechnol.* **8**: 36-7.

4. Austin G. and Morris G. (2012). Biogas production in Africa, in: R. Janssen, D. Rutz (Eds.), Bioenergy Sustain. Dev. Africa, Springer, Dordrecht, 2012.

5. Scarlat N., Dallemand J-F. and Fahl F. (2018). Biogas: Developments and perspectives in Europe. *Renew. Energ.*, **129**: 457-472.

6. Tiwary A., Williams I., Pant D. and Kishore W. (2015). Emerging perspectives on environmental burden minimisation initiatives from anaerobic digestion technologies for community scale biomass valorisation. *Renew. Sustain. Energ. Rev.*, **42**: 883-901.

7. Africa Biogas Partnership Programme, (2017). [http://africabiogas.org/](http://africabiogas.org/).

8. Mizrahi-Man O., Davenport E. and Gilad Y. (2013). Taxonomic classification of Bacterial 16S rRNA genes using short sequencing reads: Evaluation of effective study designs. *PLoS one*, **8**: e53608.

9. Campanaro S., Treu L., Kougias G., Zhu X. and Angelidaki I. (2018). Taxonomy of anaerobic digestion microbiome reveals biases associated with the applied high throughput sequencing strategies. *Scient. Rep.*, **8**: 1926.

10. Fox G., Wisotzkey J. and Jurtshuk P. (1992). How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. *Int. J. Syst. Bacteriol.*, **42**: 166-170.

11. Lacey K. and Lennon J. (2016). Scaling laws predict global microbial diversity. *Proc. Natl. Acad. Sci. USA.*, **113**: 5970-5975.

12. Lacey K. and Lennon J. (2016). Reply to Willis: Powerful predictions of biodiversity from ecological models and scaling laws. *Proc. Natl. Acad. Sci. USA.*, **113**: E5097-E5097.

13. Shoemaker W., Lacey K. and Lennon J. (2017). A macroecological theory of microbial biodiversity. *Nat. Ecol. Evol.*, **1**: 0107.

14. Dykuizen D. (1998). Santa Rosalia revisited: Why are there so many species of Bacteria? *Antonie Van Leeuwenhoek*, **73**: 25-33.

15. Schloss P., Girard R., Martin T., Edwards J. and Thrash J. (2016). Status of the Archaeal and Bacterial census: An update. *MBio.*, **7**: e00201-16.

16. Amann R. and Rosselló-Móra R. (2016). After all, only millions? *MBio.*, **7**: e000999-16.
17. Mora C., Tittensor D., Adl S., Simpson A. and Worm B. (2011). How many species are there on earth and in the Ocean? *PLOS Biol.*, 9; e1001127.
18. Wittebolle L., Marzorati M., Clement L., Balloi A., Daffonchio D., Heylen K., …, Boon N. (2009). Initial community evenness favours functionality under selective stress. *Nat.* 458: 623-6.
19. Carballa M., Regueiro L. and Lema J. (2015). Microbial management of anaerobic digestion: Exploiting the microbiome-functionality nexus. *Curr. Opin. Biotechnol.*, 33: 103-11.
20. Huse S., Huber J., Morrison H., Sogin M. and Welch D. (2007). Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol.* 8: R143.
21. Gomez-Alvarez, V., Teal, T. and Schmidt T. (2009). Systematic artifacts in metagenomes from complex microbial communities. *ISME J.*, 3: 1314-7.
22. Overbeek R., Begley T., Butler R., Choudhuri J., Diaz N., Chuang H-Y., …, Vonstein V. (2005). The Subsystems Approach to Genome Annotation and its Use in the Project to Annotate 1000 Genomes. *Nucl. Acids Res.*, 33: 17.
23. Luo G., Fotidis I. and Angelidaki I. (2016). Comparative analysis of taxonomic, functional, and metabolic patterns of microbiomes from 14 full-scale biogas reactors by metagenomic sequencing and radioisotopic analysis. *BMC Biotechnol. Biofuel*, 9: 1-12.
24. Knight, R. et al. (2012). Unlocking the potential of metagenomics through replicated experimental design. *Nat. Biotechnol. 30*: 513-520.
25. Stolze Y., Zakrzewski M., Maus I., Eikmeyer F., Jaenicke S., Rottmann N., …, Schlüter A. (2015). Comparative metagenomics of biogas-producing microbial communities from production-scale biogas plants operating under wet or dry fermentation conditions. *BMC Biotechnol. Biofuel*, 8: 14.
26. Lynd L., Weimer P., van Zyl W. and Pretorius I. (2002). Microbial cellulose utilization: Fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.*, 66: 506-577.
27. Gold D. and Martin J. (2007). Global view of the *Clostridium thermocellum* cellulosome revealed by quantitative proteomic analysis. *J. Bacteriol.*, 189: 6787-6795.
28. Seedorf H., Fricke W., Veith B., Brüggemann H., Liesegang H., Strittmatter A., …, Gottschalk G. (2008). The genome of *Clostridiumkluyveri*, a strict anaerobe with unique metabolic features. *Proc. Natl. Acad. Sci. USA*, 105: 2128-2133.
29. Kaji M., Taniguchi Y., Matsushita O., Katayama S., Miyata S., Morita S. and Okabe A. (1999). The hydA gene encoding the H₂ evolving hydrogenase of *Clostridium perfingens*: Molecular characterization and expression of the gene. *FEMS Microbiol. Lett.*, 181: 329-336.
30. Fonknechten N., Chaussonnerie S., Tricot S., Lajus A., Andreesen J., Perchat N., …, Kreimeyer A. (2010). *Clostridium sticklandii*, a specialist in amino acid degradation: Revisiting its metabolism through its genome sequence. *BMC Genomic*, 11: 555.
31. Lawson P., Song Y., Chengxu L., Denise R., Vaisanen M., Collins M. and Finegold S. (2004). *Anaerotruncus colihominis* gen. nov. sp. nov. from human feces. *IJSEM*, 54: 413-417.
32. Duncan S., Hold G., Harmsen H., Stewart C. and Flint H. (2002). Growth requirements and fermentation products of *Fusobacterium praussntizii* and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov. comb. nov. *IJSEM*, 52: 2141-2146.
33. Nokaga H., Keresztes G., Shinoda Y., Ikenaga Y., Abe M., Naito K., …, Yukawa H. (2006). Complete genome sequence of the dehalorespiring bacterium *Desulfotobacterium hafniense* YS1 and comparison with *Dehalococcoides ethenogenes*. *J. Bacteriol.*, 188: 2262-2274.
34. Debarati P., Frank W., Arick T., Bridges S., Burgess S., Yoginder S. and Lawrence M. (2010). Genome sequence of the solvent-producing bacterium *Clostridium carboxidivorans* strain P7. *J. Bacteriol.*, 192: 5554-5.
35. Rother M. and Oelgeschläger E. (2008). Carbon monoxide-dependent energy metabolism in anaerobic Bacteria. *Arch. Microbiol.*, 190: 257-269.
36. Simankova M., Chernych N., Osipov G. and Zavarzin G. (1993). *Halocella cellulolytica* gen. nov., sp. nov., a new obligately anaerobic, halophilic, cellulolytic bacterium. *Syst. Appl. Microbiol.* 16:385-389.
37. Wei C., Kunio O. and Shoichi S. (1987). Intergeneric protoplast fusion between *Fusobacterium varium* and *Enterococcus faecium* for enhancing dehydrodivanillin degradation. *Appl. Environ. Microbiol.*, 53: 542-548.
38. Guerra N., Fajardo P., Fuciños C., Amado R., Alonso E., Torrado A. and Pastrana L. (2010). Modelling the biphasic growth and product formation by Enterococcus faecium CECT 410 in realalkalized fed-batch fermentations in whey. J. Biomed. Biotechnol., 290286.
39. Enterococcus faecalis hydrogenase 3. http://biocyc.org/EFAE749511-HMP/NEWIMAGE?object=Hydrogenase-3.
40. de Bok F., Stams A., Dijkema C. and Boone D. (2001). Pathway of propionate oxidation by a syntrophic culture of Smithella propionica and Methanospirillum hungatei. Appl. Environ. Microbiol., 67: 1800-1804.
41. Schmidt A., Muller N., Schink B. and Schleheck D. (2013). A proteomic view at the biochemistry of syntrophic butyrate oxidation in Syntrophomonas wolfei. PLoS ONE, 8: e56905.
42. Dubbels B., Sayavedra-Soto L., Bottomley P. and Arp D. (2009). Thauera butanivorans sp. nov., a C2- C9 alkane-oxidizing bacterium previously referred to as ‘Pseudomonas butanovora’. Int. J. Syst. Evol. Microbiol., 59: 1577-1578.
43. Baena S., Fardeau M., Labat M., Ollivier B., Thomas P., Garcia J. and Patel B. (1998). Aminobacterium colombiense gen. nov. sp. nov., an amino acid-degrading anaerobe isolated from anaerobic sludge. Anaerobe, 4: 241-50.
44. Chertkov O., Sikorski J., Brambilla E., Lapidus A., Copeland A., Glavina Del Rio T., ..., Klenk H-P. (2010). Complete genome sequence of Aminobacterium colombiense type strain (ALA-1). Stand Genomic Sci., 2: 280-9.
45. Wei C., Kunio O. and Shiochi S. (1987). Intergeneric protoplast fusion between Fusobacterium varium and Enterococcus faecium for enhancing dehydrodivanillin degradation. Appl. Environ. Microbiol., 53: 542-54.
46. Kindaichi T., Ito T. and Okabe S. (2004). Ecophysiological interaction between nitrifying Bacteria and heterotrophic Bacteria in autotrophic nitrifying biofilms as determined by microautoradiography-fluorescence in situ hybridization. Appl. Environ. Microbiol., 70: 1641-1650.
47. Bjursel M., Martens E. and Gordon J. (2006). Functional genomic and metabolic studies of the adaptations of a prominent adult human gut symbiont, Bacteroides thetaotaomicron, to the suckling period. J. Biol. Chem., 281: 36269-36279.
48. Li H., Shen T., Wang X., Lin K., Liu Y., Lu S., ..., Du X. (2013). Biodegradation of perchloroethylene and chlorophenol co-contamination and toxic effect on activated sludge performance. Bioreour. Technol., 137: 286-93.
49. Xie G., Bruce D., Challacombe J., Chertkov O., Dettter J., Gilna P., ..., McBride M. (2007). Genomic sequence of the cellulolytic gliding bacterium, Cytophaga hutchinsonii. Appl. Environ. Microbiol., 73: 3536-3546.
50. Hiras J., Wu Y-W., Eichorst S., Simmons B. and Singer S. (2016). Refining the phylum Chlorobi by resolving the phylogeny and metabolic potential of the representative of a deeply branching, uncultivated lineage. The ISME J., 10: 833-845.
51. Wahlund T. and Madigan M. (1993). Nitrogen-fixation by the thermophilic green sulfur bacterium Chlorobium tepidum. J. Bacteriol., 175: 474-478.
52. Wahlund T. and Tabita F. (1997). The reductive tricarboxylic acid cycle of carbon dioxide assimilation: Initial studies and purification of ATP-citrate lyase from the green sulfur bacterium Chlorobium tepidum. J. Bacteriol., 179: 4859-4867.
53. Ghilarov D., Serebryakova M., Stevenson C., Hearnshaw S., Volkov D., Maxwell, A., ..., Severinov, K. (2017). The origins of specificity in the microcin-processing protease TldD/E. Structure, 25: 1549-1561.
54. Moon C., Young W., Maclean P., Cookson A. and Bermingham E. (2018). Metagenomic insights into the roles of proteoBacteria in the gastrointestinal microbiomes of healthy dogs and cats. MicrobiologyOpen, 7: e677.
55. Han K., Lim H. and Hong (1992). Acetic acid formation in Escherichia coli fermentation. Biotechnol. Bioeng., 15:663-671.
56. Textor S., Wendisch V., de Graaf A., Mu¨ller U., Linder M., Linder D. and Buckel W. (1997). Propionate oxidation in Escherichia coli: Evidence for operation of a methylcitrate cycle in Bacteria. Arch. Microbiol. 168: 428-436.
57. Hippe H., Hagelstein A., Kramer I., Swiderski J and Stackebrandt E. (1999). Phylogenetic analysis of *Formivibrio citricus*, *Propionivibrio dicarboxylicus*, *Anaerobiospirillum thomasii*, *Succinimonas amylolytica* and *Succinivibrio dextrinosolvens* and proposal of *Succinivibrionaceae fam. nov.* *Int. J. Syst. Bacteriol.*, **49**: 779-82.

58. Greetham H., Collins M., Gibson G. Giffard C., Falsen E. and Lawson P. (2004). *Sutterella stercorisicanis* sp. nov., isolated from canine faeces. *Int. J. Syst. Evol. Microbiol.*, **54**: 1581-1584.

59. Hottes AK, et al. (2004). Transcriptional profiling of *Caulobacter crescentus* during growth on complex and minimal media. *J. Bacteriol.*, **186**: 1448-1461.

60. Müller V. and Gruber G. (2003). ATP synthases: Structure, function and evolution of unique energy converters. *Cell. Mol. Life Sci.*, **60**: 474-94.

61. Manconi R. and Pronzato R. (2008). Global diversity of sponges (*Porifera: Spongillina*) in freshwater. *Hydrobiol.*, **595**: 27-33.

62. McInernery M., Mackie I. and Bryant M. (1981). Syntrophic association of butyrate degrading *Bacterium* and *Methanosarcina* enriched from bovine rumen fluid. *Appl. Environ. Microbiol.*, **42**: 826-828.

63. Meher K. and Ranade D. (1993). Isolation of propionate degrading bacterium in co-culture with a methanogen from a cattle dung biogas plant. *J. Biosci.*, **18**: 271-277.

64. Boone D. and Bryant M. (1980). Propionate-degrading bacterium, *Syntrophobacter wolini* sp. nov. gen. nov., from methanogenic ecosystems. *Appl. Environ. Microbiol.*, **40**: 626-632.

65. Mountfort D. and Bryant M. (1982). Isolation and characterization of an anaerobic syntrophic benzoate-degrading bacterium from sewage sludge. *Arch. Microbiol.*, **133**: 249-256.

66. Spring S., Sorokin D., Verborg S., Rohde M., Woyke T. and Kyrpides N. (2019). Sulfate-reducing bacteria that produce exopolymers thrive in the calcifying zone of a Hypersaline Cyanobacterial Mat. *Front. Microbiol.*, **10**:862.

67. Reichenbach H. and Dworkin M. (1991). *The myxobacteria*. In: The Prokaryotes (Starr, K.P., Stolp, H., Truper, H.G., Belows, A. and Schlegel, H.G., Eds.), Vol. 4, pp. 3415-3487. Springer, New York.

68. Dawid W. (2000). Biology and global distribution of myxobacteria in soils. *FEMS Microbiol. Rev.*, **24**: 403-427.

69. Johannessen K., Vander-Roost J., Dahl H. and Thorseth I. (2017). Biomimeralization and Fe-oxhydroxide mound formation in a diffuse hydrothermal system at the Jan Mayen vent fields. *Geochim. Cosmoch. Acta*, **202**: 101-123.

70. Vander Roost J., Thorseth I. and Dahl H. (2017). Microbial analysis of Z-proteobacteria and co-colonizers of iron mats in the Troll Wall Vent Field, Arctic Mid-Ocean Ridge. *PLOS ONE*, **12**: e0185008.

71. Hoshino T., Kuratomi T., Morono Y., Hori T., Oiwan H., Kiyokawa S. and Inagaki F. (2016). Ecophysiology of *Z-Proteobacteria* associated with shallow hydrothermal iron-oxhydroxide deposits in Nagahama Bay of Satsuma Iwo-Jima, Japan. *Front. Microbiol.*, **6**: 1554.

72. Pankratov T. and Dedysz S. (2010). *Granulicella paludicola* gen. nov., sp. nov., *Granulicella pectinivorans* sp. nov., *Granulicella aggregans* sp. nov. and *Granulicella rosea* nov., acidophilic, polymer-degrading acidobacteria from Sphagnum peat bogs. *Int. J. Syst. Evol. Microbiol.*, **60**: 2951-2959.

73. Eichorst S., Trojan D., Roux S., Herbold C., Rattei T. and Woebken D. (2018). Genomic insights into the *Acidobacteria* reveal strategies for their success in terrestrial environments. *Environ. Microbiol.*, **20**: 1041-1063.

74. Greening C., Carere C., Rushton-Green R., Harold L., Hards K., Taylor M., ..., Cook G. (2015). Persistence of the dominant soil phylum *Acidobacteria* by trace gas scavenging. *Proc. Natl. Acad. Sci. USA.*, **112**: 10497-10502.

75. Lazarev V., Levitskii S., Basovskii Y., Chukin M., Akopian T., Vereshchagin V., ..., Govorun V. (2011). Complete genome and proteome of *Acoleplasma laidlawii*. *J. Bacteriol.*, **193**: 4943-4953.

76. Freundt E., Whitcomb R., Barile M., Razin S. and Tully J. (1984). Proposal for elevation of the family *Acholeplasmataceae* to ordinal rank: *Acholeplasmatales*. *Int. J. Syst. Bacteriol.*, **34**: 346-349.
77. Ventura M., Canchaya C., Tauch A., Chandra G., FitzGerald F., Chater F. and van Sinderen D. (2007). Genomics of *Actinobacteria*: Tracing the evolutionary history of an ancient phylum. *Microbiol. Mol. Biol. Rev.*, **71**: 495-548.

78. Schell M., Karmrabüntzon M., Snel B., Vilanova D., Berger B., Pessi G., ..., Arigoni F. (2002). The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. *Proc. Natl. Acad. Sci. USA.*, **2002**: 99.

79. Zakrzewski M., Goessmann A., Jaenicke S., Jüinemann S., Eikmeyer F., Szczepanowski R., Al-Soud W., Sørensen S., Pühler A. and Schlüter A. (2012). Profiling of the metabolically active community from a production-scale biogas plant by means of high-throughput metatranscriptome sequencing. *J. Biotechnol.*, **158**: 248-58.

80. Fuerst J. and Sagulenko E. (2011). Beyond the bacterium: *Planctomycetes* challenge our concepts of microbial structure and function. *Nat. Rev. Microbiol.*, **9**: 403-413.

81. Kampmann K., Ratering S., Kramer I., Schmidt M., Zerr W. and Schnell S. (2012). Unexpected stability of *Bacteroidetes* and *Firmicutes* communities in laboratory biogas reactors fed with different defined substrates. *Appl. Environ. Microbiol.*, **78**: 2106-2119.

82. Krakat N., Westphal A., Schmidt S. and Scherer P. (2010). Anaerobic digestion of renewable biomass: Thermophilic temperature governs population dynamics of methanogens. *Appl. Environ. Microbiol.*, **77**: 1842-1850.

83. Weiss A., Jérôme V., Freitag R. and Mayer H. (2008). Diversity of the resident microbiota in a thermophilic municipal biogas plant. *Appl. Microbiol. Biotechnol.*, **81**: 163-73.

84. Maus I., Cibis K., Wibberg D., Winkler A., Stolze Y., König H., Pühler A. and Schlüter A. (2015). Complete genome sequence of the strain *Defluvitoga tunisiensis* L3, isolated from a thermophilic, production-scale biogas plant. *J. Biotechnol.*, **203**: 17-8.

85. Xia Y., Wang Y., Fang H., Jin T., Zhong H. and Zhang T. (2014). Thermophilic microbial cellulose decomposition and methanogenesis pathways recharacterized by metatranscriptomic and metagenomic analysis. *Scient. Rep.*, **4**: 6708.

86. Wang W., Xie L., Luo G., Zhou Q. and Angelidaki I. (2013). Performance and microbial community analysis of the anaerobic reactor with coke oven gas biomethanation and in situ biogas upgrading. *Bioresour. Technol.*, **146**: 23-9.

87. Siriwongrunson V., Zeng R. and Angelidaki I. (2007). Homoaacetogenesis as the alternative pathway for H2 sink during thermophilic anaerobic degradation of butyrate under suppressed methanogenesis. *Water Res.* **41**: 4204-10.

88. Xia Y., Wang Y., Fang H., Jin T., Zhong H. and Zhang T. (2014). Thermophilic microbial cellulose decomposition and methanogenesis pathways recharacterized by metatranscriptomic and metagenomic analysis. *Scient. Rep.*, **4**: 6708.

89. Eichorst S. and Kuske C. (2012). Identification of cellulose-responsive bacterial and fungal communities in geographically and edaphically different soils by using stable isotope probing. *Appl. Environ. Microbiol.*, **78**: 2316-2327.

90. Appel J. and Schulz R. (1998). Hydrogen metabolism in organisms with oxygenic photosynthesis: Hydrogenases as important regulatory devices for a proper redox poising? *J. Photochem. Photobiol. Ser. B*, **47**: 1-11.

91. Bergman B., Gallon J., Rai A. and Stal L. (1997). N2 fixation by non-heterocystous *Cyanobacteria*. *FEMS Microbiol. Rev.*, **19**: 139-185.

92. Lindblad P. (1999). *Cyanobacterial* H2-metabolism: Knowledge and potential/strategies for a photo-biotechnological production of H2. *Biotecnol. Appl.* **16**: 141-144.

93. Howarth D. and Codd A. (1985). The uptake and production of molecular hydrogen by unicellular *Cyanobacteria*. *J. Gen. Microbiol.*, **131**: 1561-1569.

94. Weisshaar H. and Böger P. (1985). Pathways of hydrogen uptake in the *Cyanobacterium Nostoc muscorum*. *Arch. Microbiol.*, **142**: 349-353.

95. Siege A., Kamke J., Hochmuth T., Piel J., Richter M., Liang C., Dandekar T. and Hentschel U. (2011). Single-cell genomics reveals the lifestyle of *Poribacteria*, a candidate phylum symbiotically associated with marine sponges. *The ISME J.*, **5**: 61-70.
96. Pelletier E., Kreimeyer A., Bocs S., Rouy Z., Gyapay G., Chouari R., ..., Denis Le Paslier (2008). “Candidatus Cloacamonas acidaminovorans”: Genome sequence reconstruction provides a first glimpse of a new Bacterial division. J. Bacteriol., 190: 2572-2579.
97. Guivarch E., Pons A., Christophe G., Creuly C. and Dussap C-G. (2010). Analysis of a continuous culture of Fibrobacter succinogenes S85 on a standardized glucose medium. Bioprocess Biosyst. Eng. 33: 417-25
98. Ye Q., Roh Y., Carroll S., Blair B., Zhou J., Zhang C. and Fields M. (2004). Alkaline anaerobic respiration: Isolation and characterization of a novel alkaliphilic and metal-reducing bacterium. Appl. Environ. Microbiol., 70: 5595-5602.
99. Fisher E., Dawson A., Polshyna G., Lisak J., Crable B., Detter J., ..., Fields M. (2008). Transformation of inorganic and organic arsenic by Alkaliphilus oremlandii sp. nov. strain OhILAs. Ann. NY. Acad. Sci. 1125: 230-241.
100. Daims H., Lebedeva E., Pjevac P., Han P., Herbold C., Albertsen M., ..., Wagner M. (2015). Complete nitrification by Nitrospira bacteria. Nat., 528: 504-509.
101. Koch H., van Kessel M. and Lücke S. (2019). Complete nitrification: Insights into the ecophysiology of comammox Nitrospira. Appl. Microbiol. Biotechnol., 103:177-189
102. Kotsyurbenko O., Glagolev M., Nozhevnikova A. and Conrad R. (2001). Competition between haloarchaeal Archaea and methanogenic Archaea for hydrogen at low temperature. FEMS Microbiol. Ecol., 38: 153-159.
103. Thauer R., Kaster A., Seedorf H., Buckel W. and Hedderich R. (2008). Methanogenic Archaea: Ecological relevant differences in energy conservation. Nat. Rev. Microbiol., 6: 579-589.
104. Deppenmeier U. and Müller V. (2008). Life close to the thermodynamic limit: How methanogenic Archaea conserve energy. Result and Problem Cell Different., 45:123-152.
105. Rother M. and Melchaf W. (2004). Anaerobic growth of Methanosarcina acetivorans C2A on carbon monoxide: An unusual way of life for a methanogenic archaeon. Proc. Natl Acad. Sci.U.SA, 101: 16929-16934.
106. Anderson I., Siepawska-Lupa M., Goltsman E., Lapidus A., Copeland A., Glavina Del Rio T., Kyrpides N. (2009). Complete genome sequence of Methanocorpusculum labreanum type strain Z. Stand Genomic Sci., 1:197-203.
107. Gilmore S., Henske J., Sexton J., Solomon K., Seppälä S., Yoo J., ..., O’Malley M. (2017). Genomic analysis of methanogenic Archaea reveals a shift towards energy conservation. BMC Genomic, 18: 639.
108. Zellner G., Stackebrandt E., Messner P., Tindall B., Demacario E., Kneifel H., Sleytr U. and Winter J. (1989). Methanocorpusculaceae fam-nov, represented by Methanocorpusculum parvum, Methanocorpusculum sinense spec-nov and Methanocorpusculum bavarium spec-nov. Arch. Microbiol., 151: 381-90.
109. Anderson I., Siepawska-Lupa M., Lapidus A., Nolan M., Copeland A., Glavina Del Rio T., ..., Kyrpides N. (2009). Complete genome sequence of Methanothellicus marisnigri Romesser et al. 1981 type strain JR1. Stand.Genomic Sci., 1: 189-196.
110. Rivera M. and Lake J. (1996). The phylogeny of Methanopyrus kandleri. Int. J.Syst. Bacteriol., 46: 348-351.
111. Burggraf S., Fricke H., Neuner A., Kristjansson J., Rouvier P., Mandelco L., ..., Stetter K. (1990). Methanococcus igneus sp. nov., a novel hyperthermophilic methanogen from a shallow submarine hydrothermal system. Syst. Appl. Microbiol., 13: 263-269.
112. Ding J., Chen S., Lai M. and Liao T. (2017). Haloterrigena mahii sp. nov., an extremely halophilic archaeon from a solar saltern. Int. J. Syst. and Evolut. Microbiol., 67: 1333-1338.
113. Oren A. (2006). The order HaloBacteriales. In Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E, editors. The Prokaryotes. A Handbook on the Biology of Bacteria. Volume 3: Archaea. Bacteria: Firmicutes, Actinomycetes. Springer; pp. 113-164.
114. Huber H. and Stetter K. (2006). Thermoplasmatales. In Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E, editors. The Prokaryotes. A Handbook on the Biology of Bacteria. Volume 3: Archaea. Bacteria: Firmicutes, Actinomycetes. Springer; pp. 101-112.
115. Lang, K., Schuldes, J., Kling, A., Poehlein, A., Daniel R., and Brunea A. (2015). New mode of energy metabolism in the seventh order of methanogens as revealed by comparative genome analysis of “Candidatus Methanoplasma termitium”. Appl. Environ. Microbiol., 81: 1338-52.

116. Gorlas A., Alain K., Bienvenu N. and Geslin C. (2013). Thermococcus prieurii sp. nov., a hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent. Int. J. Syst. Evol. Microbiol., 63: 2920e2926.

117. Cossu M., Da-Cunha V., Toffano-Nioche C., Forterre P. and Oberto J. (2015). Comparative genomics reveals conserved positioning of essential genomic clusters in highly rearranged Thermococcales chromosomes. Biochimie, 118: 313e321.

118. Brasen C., Esser D., Rauch B. and Siebers B. (2014). Carbohydrate metabolism in Archaea: Current insights into unusual enzymes and pathways and their regulation. Microbiol. Mol. Biol. Rev., 78: 89e175.

119. Wang X., Gao Z., Xu X. and Ruan L. (2011). Complete genome sequence of Thermococcus sp. Strain 4557, a hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent area. J. Bacteriol., 193: 5544-5545

120. Lee H., Kang S., Bae S., Lim J., Cho Y., Kim Y., …, Lee J-H. (2008). The complete genome sequence of Thermococcus onnurineus NA1 reveals a mixed heterotrophic and carboxydrotrophic metabolism. J. Bacteriol., 190: 7491-7499

121. Hartzell P. and Reed D. (2006). The genus Archaeoglobus. In Dworkin M, Falkow S, Rosenberg E., Schleifer K., Stackebrandt E., editors. The Prokaryotes. A Handbook on the Biology of Bacteria. Volume 3: Archaea. Bacteria: Firmicutes, Actinomycetes. Springer, pp. 82-100.

122. Kashefi K., Tor J., Holmes D., Gaw Van Praagh C., Reysenbach A-L and Lovley D. (2002). Geoglobus ahangari gen. nov., sp. nov., a novel hyperthermophilic archaeon capable of oxidizing organic acids and growing autotrophically on hydrogen with Fe(III) serving as the sole electron acceptor. Int. J. Syst. Evol. Microbiol., 52: 719-28.

123. Dai X., Wang H., Zhang Z., Li K., Zhang X., Mora-López M., …, Huang L. (2016). Genome sequencing of Sulfolobus sp. A20 from Costa Rica and comparative analyses of the putative pathways of carbon, nitrogen, and sulfur metabolism in various sulfolobus strains. Front. Microbiol., 7: 1902.

124. Sakai H. and Kurosawa N. (2019). Complete genome sequence of the Sulfodiicoccus acidiphilus strain HS-1T, the first crenarchaeon that lacks polB3, isolated from an acidic hot spring in Ohwaku-dani, Hakone, Japan. BMC Res. Notes, 12: 444.

125. Berg I., Kockelkorn D., Buckel W and Fuchs G. (2007). A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. Sci. 318: 1782-1786.

126. Siebers B., Tjaden B., Michalke K., Do’rr C., Ahmed H., Zaparty M., …, Hensel R. (2004). Reconstruction of the central carbohydrate metabolism of Thermoproteus tenax by use of genomic and biochemical data. J. Bacteriol., 186: 2179-2194.

127. Zaparty M., Tjaden B., Hensel R. and Siebers B. (2008). The central carbohydrate metabolism of the hyperthermophilic crenarchaeote Thermoproteus tenax: Pathways and insights into their regulation. Arch. Microbiol., 190: 231-245.

128. Cossu M., Cunha V., Toffano-Nioche C., Forterre P. and Oberto J. (2015). Comparative genomics reveals conserved positioning of essential genomic clusters in highly rearranged Thermococcales chromosomes. Biochimie, 118: 313e321.

129. Prokofeva M., Kostrikina N., Kolganova T., Tourova T., Lysenko A., Lebedinsky A., and Bonch-Osmolovskaya E. (2009). Isolation of the anaerobic thermoacidophilic crenarchaeote Acidilobus saccharovorans sp. nov. and proposal of Acidilobales ord. nov., including Acidilobaceae fam. nov. and Caldiphaeraceae fam. nov. Int. J. Syst. Evol. Microbiol., 59: 3116-3122.

130. Mardanov A., Svetlitchnyi V., Beletsky A., Prokofeva M., Bonch-Osmolovskaya, E., Ravin N. and Skryabin K. (2009). The genome sequence of the Crenarchaeon Acidilobus saccharovorans supports a New Order, Acidilobales, and suggests an important ecological role in terrestrial acidic hot springs. Appl. Environ. Microbiol., 77: 5652-5657.
131. Hallam S., Minner T., Schleper C., Preston C., Roberts K., Richardson P. and DeLong E. (2006). Pathways of carbon assimilation and ammonia oxidation suggested by environmental genomic analyses of marine Crenarchaeota. PLOS Biol. 4: e95.

132. Hallam S., Konstantinidis K., Putnam N., Schleper C., Watanabe Y., Sugahara J., …, DeLong E. (2006). Genomic analysis of the uncultivated marine Crenarchaeote *Cenarchaeum symbiosum*. *Proc. Natl. Acad. Sci. USA*, 103: 18296-18301.

133. Walker C., de la Torre J., Klotz M., Urakawa H., Pinel N., Arp D., …, Stahl D. (2010). *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine *Crenarchaeae*. *Proc. Natl Acad. Sci.*, 107: 8818-8823.

134. Hu’gler M., Huber H., Stetter K. and Fuchs G. (2003). Autotrophic CO2 fixation pathways in *Archea (Crenarchaeota)*. *Arch. Microbiol.*, 179: 160-173.

135. Alonso-Sáez L., Waller A., Mende D., Bakker K., Farnelid H., Yager PL., …, Bertilsson S. (2012). Role for urea in nitrification by polar marine *Archea. Proc. Natl. Acad. Sci. USA*, 109: 17989-17994.

136. Podar M., Podar M., Makarova K., Graham D., Wolf Y., Koonin E. and Reysenbach A-L. (2013). Insights into *archaeal* evolution and symbiosis from the genomes of a Nanoarchaeon and its inferred Crenarchaeal host from Obsidian Pool, Yellowstone national park. *Biol. Direct.*, 8: 9.

137. Munson-McGee J., Field E., Bateson M., Rooney C., Stepauenas R. and Young M. (2015). *Nanoarchaeota*, their *sulfolobales* host, and *Nanoarchaeota* virus distribution across Yellowstone national park hot springs. *Appl. Environ. Microbiol.*, 81: 7860-7868.

138. Wurch L., Giannone R., Belisle B., Swift C., Utturkar S., Hettich R., Reysenbach A-L. and Podar M. (2016). Genomics-informed isolation and characterization of a symbiotic *Nanoarchaeota* system from a terrestrial geothermal environment. *Nat. Commun.*, 7: 1-10.

139. Pinto O., Silva T., Vizzotto C., Santana R., Lopes F., Silva B., Thompson F. and Kruger R. (2020). Genome-resolved metagenomics analysis provides insights into the ecological role of *Thaumarchaeota* in the Amazon river and its plume. *MBMC Microbiol.*, 20: 13.

140. Bernailler A., Fonty G., Bonnemoy F. and Gouet P. (1992). Degradation and fermentation of cellulose by the rumen anaerobic fungi in axenic cultures or in association with cellulolytic bacteria. *Curr. Microbiol.*, 25: 143-148.

141. Costa V., Amorim M., Reis E., Quintanilha A. and Moradas-Ferreira P. (1997). Mitochondrial superoxide dismutase is essential for ethanol tolerance of *Saccharomyces cerevisiae* in the post-diauxic phase. *Microbiol.*, 143: 1649-1656.

142. Viigand K., Tammus K. and Alamäe T. (2005). Clustering of MAL genes in *Hansenula polymorpha*: Cloning of the maltose permease gene and expression from the divergent intergenic region between the maltose permease and maltase genes. *FEMS Yeast Res.*, 5: 1019-1028.

143. Yurimoto H., Oku M. and Sakai Y. (2011). Yeast methylotrophy: Metabolism, Gene regulation and peroxisome homeostasis. *Int. J. Microbiol.*, 2011: 101298.

144. McHardy AC., Fan L., and Yu X. (2017). Comparative whole-genome analysis reveals artificial selection effects on *Ustilago scleuenata* genome. *DNA Rex.*, 24: 635-648.

145. Ren X., Liu Y., Tan Y., Huang Y., Liu Z. and Jiang X. (2019). Sequencing and functional annotation of the whole genome of *Shiraia bambusicola*. *Genes, Genome, Genetics*, 10: 23-35.

146. Urquhart A., Mondo S., Mäkelä M., Hane J., Wiebenga A., He G., …, Idnurm A. (2018). Genomic and genetic insights into a cosmopolitan fungus, *Paecilomyces variotii* (*Eurotiales*). *Front. Microbiol.*, 9: 3058.

147. Ohm R., Feau N., Henrissat B., Schoch C., Horwitz B., Barry K., …, Grigoriev I. (2012). Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen *Dothideomycetes* fungi. *PLOS Pathog.*, 8: e1003037.

148. Li J., Yu L., Yang J., Dong L., Tian B., Yu Z., …, Zhang K. (2010). New insights into the evolution of subtilisin-like serine protease genes in *Pezizomycotina*. *BMC Evol. Biol.*, 10: 68.

149. Murat C., Payen T., Noel B., Kuo A., Morin E., Chen J., Kohler A. Krizsán K., …, Martin F. (2018). *Pezizomycetes* genomes reveal the molecular basis of ectomycorrhizal truffle lifestyle. *Nat. Ecol. Evol.*, 2: 1956-1965.
150. Larriba E., Martín-Nieto J. and Lopez-Llorca L. (2012). Gene cloning, molecular modeling, and phylogenetics of serine protease P32 and serine carboxypeptidase SCP1 from nematophagous fungi Pochonia rubescens and Pochonia chlamydospora. Can. J. Microbiol., 58: 815-27.

151. Voigt C., Schäfer W. and Salomon S. (2005). A secreted lipase of Fusarium graminearum is a virulence factor required for infection of cereals. The Plant J., 42: 364-375.

152. Olivieri F., Eugenia Zanetti M., Oliva C., Covarrubias A. and Casalongue C. (2002). Characterization of an extracellular serine protease of Fusarium eumartii and its action on pathogenesis related proteins. Eur. J. Plant. Pathol., 108: 63-72.

153. Agrawa Y., Khatri I., Subramanian S. and Shenoy B. (2015). Genome sequence, comparative analysis and evolutionary insights into chitinases of entomopathogenic fungus, Hirsutella thompsonii. Genome Biol. Evol., 9: 916-930.

154. Yang Y., Liu B., Du X., Li P., Liang B., Cheng X., ..., Wang S. (2014). Complete genome sequence and transcriptomics analyses reveal pigment biosynthesis and regulatory mechanisms in an industrial strain, Monascus purpureus YY-1. Sci. Rep., 5: 8331.

155. Galagan J., Calvo S., Borkovich K., Selker E., Read N., Jaffe D., ..., Birren B. (2003). The genome sequence of the filamentous fungus Neurospora crassa. Nat., 422: 859-68.

156. Zheng H., Zhong Z., Shi M., Zhang L., Lin L., Hong Y., ..., Wang Z. (2018). Comparative genomic analysis revealed rapid differentiation in the pathogenicity related gene repertoires between Pyricularia oryzae and Pyricularia pennisetii isolated from a Pennisetum grass. BMC Genomes, 19: 927.

157. Zhang N., Cai G., Price D., Crouch J., Gladieux P., Hillman B., ..., Bhattacharya D. (2018). Genome wide analysis of the transition to pathogenic lifestyles in Magnaportheales fungi. Sci. Rep., 8: 5862.

158. Bugos R., Sutherland J. and Adler J. (1988). Phenolic compound utilization by the soft rot fungus Lecythophora hoffmannii. Appl. Environ. Microbiol., 54: 1882-1885.

159. Leonhardt S., Büttner E., Gebauer A., Hofrichter M. and Kellner H. (2018). Draft genome sequence of the sordariomycete Lecythophora (Coniochaeta) hoffmannii CBS 245.38. Genome Announc., 6: e01510-17.

160. Lindahl B., de Boer W. and Finlay R. (2010). Disruption of root carbon transport into forest humus stimulates fungal opportunists at the expense of mycorrhizal fungi. ISME J., 4: 872-881.

161. Wang Q.-M., Theelen B., Groenewald M., Bai F.-Y. and Boekhout T. (2014). Moniliellomycetes and Malasseziomycetes, two new classes in Ustilaginomycotina. Persoonia, 33: 41-47.

162. Sánchez C. (2009) Lignocellulosic residues: Biodegradation and bioconversion by fungi. Biotechnol. Advances, 27: 185-194.

163. James T., Letcher P., Longcore J., Mozley-Standridge S., Porter D., Powell M., Griffith G. and Vilgalys R. (2006). A molecular phylogeny of the flagellated fungi (Chytriomyctota) and description of a new phylum (Blastocladiomycota). Mycologia, 98: 860-871.

164. Bart T., van de Vossenberg L., Warris S., Nguyen H., van Gent-Pelzer M., Joly D., ..., van der Lee T. (2019). Comparative genomics of chytrid fungi reveal insights into the obligate biotrophic and pathogenic lifestyle of Synchytrium endobioticum. Sci. Rep., 9:8672.

165. Karov S., Mamanazarova K., Popova O., Aleoshin V., James T., Mamkaeva M., ..., Vishnyakov A. (2017). Monoblepharidomycetes diversity includes new parasitic and saprotrophic species with highly intronized rDNA. Fungal Biol., 121: 729-741.

166. Van der Giezen M. (2009). Hydrogenosomes and mitosomes: Conservation and evolution of functions. J. Euk. Microbiol., 56: 221-231.

167. Embley T., van der Giezen M., Horner D., Dyal P. and Foster P. (2002). Mitochondria and hydrogenosomes are two forms of the same fundamental organelle. Phil. Trans. R. Soc. Lond. B. 358: 191-203.

168. Hibbett D., Binder M., Bischoff J., Blackwell M., Cannon P., Eriksson O., ..., Zhang N. (2007). A higher-level phylogenetic classification of the Fungi. Mycolog. Res., 111: 509-547.

169. Hoffmann K., Voigt K. and Kirk P. (2011). Mortierellomycotina subphyl. nov., based on multigene genealogies. Mycotoxan, 115: 353-363.
onse to a pH perturbation of a cellulose...

170. Wagner L., Stielow B., Hoffmann K., Petkovits T., Papp T., Vágvölgyi C., ... Voigt K. (2013). A comprehensive molecular phylogeny of the Mortierellales (Mortierellomycotina) based on nuclear ribosomal DNA. Persoonia, 30: 77-93.

171. Cheng Y., Edwards J., Allison G., Zhu W. and Theodorou M. (2009). Diversity and activity of enriched ruminal cultures of anaerobic fungi and methanogens grown together on lignocellulose in consecutive batch culture. Bioresource Technol., 100: 4821-4828.

172. Nakashimada Y., Srinivasan K., Murakami M. and Nishio N. (2000). Direct conversion of cellulose to methane by anaerobic fungus Neocallimastix frontalis and defined methanogens. Biotechnol. Lett., 22: 223-227.

173. Lee S., Ha J. and Cheng K-J. (2000). Relative contributions of Bacteria, Protozoa, and fungi to in vitro degradation of orchard grass cell walls and their interactions. Appl. Environ. Microbiol., 66: 3807-3813.

174. Leis S., Dresch P., Peintner U., Fliegerová K., Sandbichl A., Insam H. and Podmirseg S. (2014). Finding a robust strain for biomethanation: Anaerobic fungi (Neocallimastigomycota) from the Alpine ibex (Capra ibex) and their associated methanogens. Anaerobe, 29: 34-43.

175. Zhou J., Deng Ye., Zhang P., Xue K., Lianga. Y., Van Nostrand J. ..., Arkin A. (2014). Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. Proc. Natl Acad. Sci. USA, E836-E845.

176. Ofiteru I., Lunnc M., Curtisa T., Wellsd G., Criddled C., Francise C., and Sloan W. (2010). Combined niche and neutral effects in a microbial wastewater treatment community. Proc. Natl Acad. Sci. USA, 107: 15345-15350.

177. Grant T. (2010). The Randomized Complete Block Design (RCBD). Department of Horticulture and Crop Science, OARDC, The Ohio State University, USA.

178. Boaro A., Kim Y., Konopka A., Callister S. and Ahring B. (2014). Integrated ‘Omnics’ analysis for studying the microbial community response to a pH perturbation of a cellulose-degrading bioreactor culture. FEMS Microbiol. Ecol., 90: 802-815.

179. Menon, S. (2014). ArcGIS 10.3: The Next Generation of GIS Is Here. Accessed June, 2019.

180. Oksenberg L., Cannell, C. and Kalton G. (1991). New strategies for pretesting survey questions. J. Off. Statist., 7: 349-365.

181. Campanaro, S., Treu L., Kougias P., Francisci D., Valle G. and Angelidaki I. (2016). Metagenomic analysis and functional characterization of the biogas microbiome using high throughput shotgun sequencing and a novel binning strategy. BMC Biotechnol. Biofuel, 9: 26.

182. Caporaso J., Lauber C., Walters W., Berg-Lyons D., Huntley J., Fierer N., ... Knight R. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 6: 1621-4.

183. Zhou Q., Su X., and Ning K. (2014). Assessment of quality control approaches for metagenomic data analysis. Scient. Rep. 4: 6957.

184. Bolger A., Lohse M. and Usadel B. (2014). Trimmomatic: A flexible trimmer for 588 Illumina sequence data. BMC Bioinformatics, 30: 2114-2120.

185. Nurk S., Meleshko D., Korobeynikov A., and Pevzner P. (2017). metaSPAdes: A new versatile metagenomic assembler. Genome Res., 27: 824-834.

186. Cox M., Peterson D. and Biggs P. (2010). SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinformatics, 11: 485.

187. Keegan K., Trimble W., Wilkening J., Wilke A., Harrison T., D’Souza M. and Meyer F. (2012). A platform-independent method for detecting errors in metagenomic sequencing data: DRISEE. PLOS Comput. Biol., 8(6):e1002541.

188. Langmead B., Trapnell C., Pop M. and Salzberg S. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol, 10: R25.

189. Wilke A., Harrison T., Wilkening J., Field D., Glass E., Kyrpides N., ..., Meyer F. (2012). The M5nr: A novel non-redundant database containing protein sequences and annotations from multiple sources and associated tools. BMC Bioinformatic, 13: 141.

190. Rho M., Tang H. and Ye Y. (2010). FragGeneScan: Predicting genes in short and error-prone reads. Nucl. Acids Res., 38: e191

191. Kent W. (2002). BLAT-the BLAST-like alignment tool. Genome Res., 12: 656-664.
192. Field D., Sterk P., Kottmann R., Wim De Smet J., Amaral-Zettler L., Cochrane G., …, et al. (2014). Genomic standards consortium projects. *Stand. Genome Sci.*, 9: 599-601.

193. Glass E., Dribinsky Y., Yilmaz P., Levin H., Van Pelt R., Wendel D., …, Schriml L. (2014). MIxS-BE: A MIxS extension defining a minimum information standard for sequence data from the built environment. *ISME J.*, 8: 1-3.

194. Angiuoli S., Matalka M., Gussman A., Galens K., Vangala M., Riley D., …, Fricke W. (2011). CloVR: A virtual machine for automated and portable sequence analysis from the desktop using cloud computing. *BMC Bioinformatic*, 12: 356.

195. Terry Speed (2003). *Statistical Analysis of Gene Expression Microarray Data*. Chapman and Hall/CRC, 2003.

196. Elmore K. and Richman M. (2000). Euclidean Distance as a Similarity Metric for Principal Component Analysis. *Amer. Meteorol. Soc.*, 129: 540-549.

197. Meyer F., Paarmann D., D'Souza M., Olson1 R., Glass E., Kubal M., …, Edwards R. (2008). The metagenomics RAST server-a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatic*, 9: 386.

198. Ondov B., Bergman N. and Phillippy A. (2011). Interactive metagenomic visualization in a Web browser. *BMC Bioinformatic*, 12: 385.