npj | clean water

www.nature.com/npjcleanwater

ARTICLE OPEN

Check for updates

# Characterization of the core microbial community governing acidogenic processes for the production of valuable bioproducts

Qidong Yin[1]✉, Guangxue Wu[1] and Piet N. L. Lens[1]

Volatile fatty acids (VFAs) and alcohols generated from acidogenic processes are valuable bioresources. However, how the diversity of acidogenic microorganisms and environmental factors affect their generation are still poorly understood. In this study, 18 different inocula and 42 sludges from acidogenic lab-scale reactors were collected to analyze the microbial communities and their metabolic potential using 16S rRNA genes high throughput sequencing coupled with PICRUSt2. 23 out of 30732 distinctive amplicon sequence variants were identified as the core features and 34.8% of them (e.g., *Clostridium spp.*) were positively correlated with the generation of the most common product acetate. PICRUSt2 shows that an average of 27% of predicted fermentation-pathway genes was assigned to the core features, suggesting their crucial roles in acidogenesis. From the network aspect, the acidogenic network had a slightly higher number of nodes (12%), but significantly lower numbers of edges (109%) and neighbors (132%) compared with the inoculum network. A total of 28 independent subnetworks from large to small scales were extracted from the acidogenic network. The decentralized distribution of core features in these subnetworks emphasized their non-co-occurring relationships. The electrode potential was the most significant environmental variable (48.2–49.3% of the explanation), positively affecting the distribution of more than 50% of the core features and fermentation pathways. Results of this study emphasized the importance of core features rather than microbial diversity in acidogenic performance and highlighted the response of the core microbial community to environmental changes, which may be applied in practical applications to optimize acidogenic performance.

*npj Clean Water* (2022)5:39 ; https://doi.org/10.1038/s41545-022-00180-3

## INTRODUCTION

Volatile fatty acids (VFAs) and alcohols, e.g., acetate, propionate, butyrate, and ethanol, have wide industrial applications but their current production mainly relies on fossil sources[1]. Acidogenic processes, e.g., homoacetogenesis, solventogenesis, fermentation, and acidification, can be alternative environmentally friendly approaches for producing VFAs and alcohols. In these processes, microorganisms catalyze $CO_2$ or other organic wastes into VFAs and alcohols with complex interactions and metabolic pathways, simultaneously reducing $CO_2$ emissions and achieving contaminant removal.

Understanding the distribution pattern of microbial communities and identification of functional microbes are key to improving the system performance of acidogenic processes. Currently, many studies have provided novel insights into the diversity, distribution, and functions of microbial populations in environmental biotechnologies through bioinformatic analysis. For example, Wu et al.[2] examined the global diversity and biogeography of microbial populations in activated sludge collected from wastewater treatment plants and found that the bacterial composition was driven by stochastic processes. Among the high diversity of the microbial communities, 28 operational taxonomic units were identified as a global core microbial community which could be strongly linked to system performance[2]. Microbial community composition and diversity in anaerobic digestion systems were also investigated and could be linked to renewable biogas production[3,4].

Despite the rapid advancement of biorefinery technologies, the microbial diversity and ecology of acidogenic systems are less understood. Some studies focused on the effects of one or two specific environmental variables on the performance of acidogenic systems. For example, Zhang et al.[5] investigated the effects of fermentation substrate and zero-valent iron on the microbial community and metabolic function in a fermentation system. Park et al.[6] investigated the metabolic flux and potential function of the microbial community in an acidogenic dynamic membrane bioreactor. Nevertheless, microbial profiles of acidogenic systems such as microbial diversity and the identification of the core microbial community are still far from being comprehensively characterized.

The microbial composition in acidogenic systems may be affected by many environmental factors such as pH, temperature, inoculum, and substrate type[5,7,8]. These environmental factors provide different microbial niches affecting the microbial community dynamics and metabolic pathways. For instance, solventogenesis usually occurs at a pH lower than 5.5[9], while propionate-producing bacteria prefer to grow at a neutral pH[7]. Besides, microbial electrosynthesis has been recently proposed as an efficient approach to control and stabilize fermentation processes[10]. Microbial electrosynthesis has been extensively used to produce different types of soluble molecules (e.g., alcohols or carboxylic acids) from the reduction of $CO_2$ or the fermentation of organic matter, and to overcome some limitations of metabolic reactions[11,12]. Such bioelectrochemical systems exert significant advantages on microbial metabolism, interspecies interactions,

[1]National University of Ireland Galway, Galway, Ireland. ✉email: qidongyin@qq.com

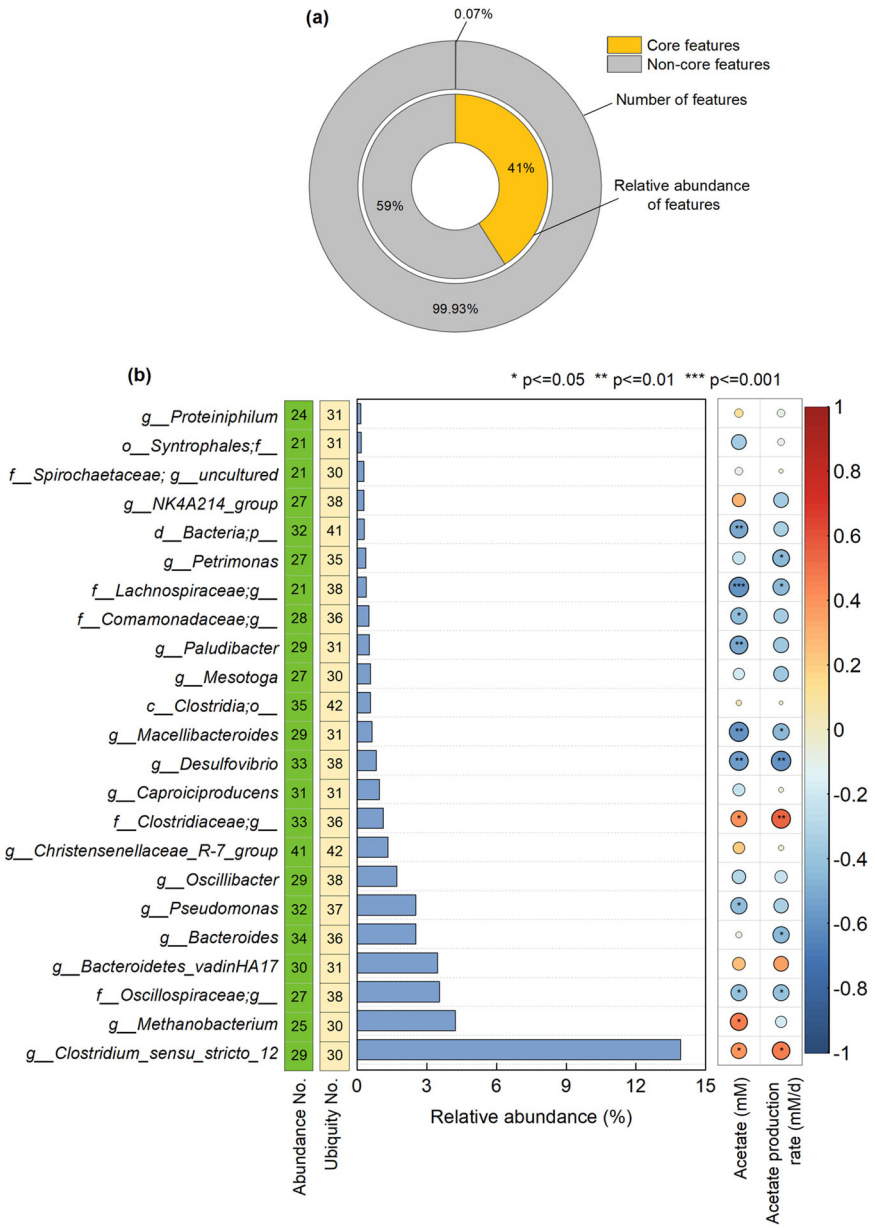


**Fig. 1 Characterizations of core features identified. a** Percentage and relative abundance of the core features and **b** the taxonomic composition of core features at the genus level and their frequency of abundance (green column), frequency of ubiquity (yellow column), and correlation with acetogenic performances.

and the selection of microbial populations[10]. Nonetheless, the detailed effects of electrode potential remain largely unrevealed. Accordingly, comprehensive profiling of microbial communities in acidogenic systems with different environmental variables may provide valuable information for predicting microbial ecosystems responding to environmental changes and benefit the optimization of biorefinery technologies.

In this study, 18 different inocula and 42 samples collected from lab-scale acidification reactors with different environmental variables were analyzed to explore the microbial community and diversity of acidogenic systems. The core microbial community and their metabolic pathways in acidogenic samples were identified. The acidogenic microbial network and subnetworks were established to highlight the distribution and topological structures of core features from the network aspect. Furthermore, the effects of key environmental variables, e.g., pH, substrate type, temperature, inoculum, and bioelectrochemical potential on the core microbial community and metabolic pathways were analyzed. Statistical analysis was conducted to investigate the potential correlation between the acetogenic performance and core microbial community or predicted functional pathways. The distribution pattern and metabolic pathways of core features revealed in this study can help guide the regulation of core microbial community to optimize acidogenic practices.

## RESULTS AND DISCUSSION

### Core microbial community

A total of 3,107,841 sequences were obtained from all 60 samples, with an average of 51,797 sequences per sample. The core microbial community was determined by the abundance and occurrence frequency of ASVs[2]. As shown in Fig. 1a, 23 out of 30732 ASVs were identified as the core features that accounted for

41% of the sequences in all acidogenic samples investigated. Most of the core features are acidogenic bacteria and are distinct from the core microbial community identified in domestic wastewater treatment plants or anaerobic digestion processes producing methane[2,3], suggesting that the acidogenic system selects a unique core microbial community.

At the phylum level, nine core features belonged to *Firmicutes*, followed by *Bacteroidota* (6), i.e., 65.2% of the core features belonged to these two phyla. At the genus level, *Clostridium sensu stricto* 12 was identified as the most abundant core feature, accounting for 13.9% of the sequence abundance. *Clostridium* is a well-known acetogen widely existing in acidogenic processes[13–15]. Surprisingly, the second most abundant core feature (4.2% of the sequence abundance) was a typical $H_2$ utilizing methanogen *Methanobacterium*. *Methanobacterium* was ubiquitously detected in 71.4% and abundant in 59% of all acidogenic samples. One explanation is that many inoculum samples were collected from anaerobic digesters producing methane, and although bromoethane sulfonate (BES) was added into some of these samples, it seemed to fail to completely eliminate *Methanobacterium*[9,16]. Besides, it is reported that hydrogenotrophic methanogens have a higher $H_2$ competitiveness than homoacetogens at a high $H_2$ partial pressure[17]. *Desulfovibrio* was also identified as a core feature, reflecting the importance of incomplete oxidizing sulfate-reducing bacteria in acidogenic systems that develop out of inocula from full-scale anaerobic digesters.

The prevailing microorganisms are often suggested to be responsible for maintaining system performance[3]. To evaluate the correlation between the core features and acidogenic system performance, the highest acetate concentration and the maximum acetate-production rate were used to characterize acetogenic performance since acetate was the most common product in all samples. Three core features were positively correlated with both the acetate concentration and acetate-production rate (Fig. 1b), including *Clostridium sensu stricto* 12 ($p \le 0.05$), *Bacteroidetes* vadinHA17, and one ASV belonged to the phylum *Clostridiaceae* ($p \le 0.05$). The other 5 core features were positively correlated with at least one acetogenic performance while the remaining core features were negatively correlated with acetogenic performance. The core features that were negatively associated with acetate production might be critical for maintaining other acidogenic functions such as propionate and butyrate production. For example, butyrate can be produced via bioelectrochemical chain elongation from $CO_2$ and acetate, or via acetyl-CoA reduction with $H_2$ as electron donor[18]. All these processes consume acetate or precursors of acetate production. Consistent with this, the butyrate producers *Oscillibacter*[19] and *Caproiciproducens*[20] were negatively correlated with acetogenic performance. On the other hand, core features positively correlated with the acetogenic function might be the winners of substrate competition since acetate is the end product of homoacetogenesis, fermentation, and acidification. The correlation result shown in Supplementary Fig. 1 may support this hypothesis because many correlations among core features were negative. Particularly, *Clostridium sensu stricto* 12 and the ASV belonging to the phylum *Clostridiaceae* were negatively correlated with most (73–77%) of the other core features.

## Microbial network

Microbial correlations were also presented by the network analysis (Fig. 2). As shown in Fig. 2a, a complex microbial network (855 genera involved) was observed with more than 15 thousand correlations (Pearson correlation coefficient >0.7, $p \le 0.05$) existing in the acidogenic community. The topological parameters of the acidogenic network are listed in Supplementary Table 1. The low network density (0.04) and high heterogeneity (0.9) suggested that the majority of genera tended to have few connections[21]. Besides, the low network centralization (0.1) implied that the centralities of genera are relatively similar[22]. Furthermore, the whole network consisted of many high-density circle-type subnetworks. There might be multiple relatively independent ecosystems within the acidogenic system investigated.

To investigate the in-depth structure of the acidogenic network, subnetworks were further extracted. A total of 28 subnetworks were formed in the acidogenic network. These subnetworks were marked from 1 to 28 according to the number of nodes. Figure 2b shows the structures of the largest 12 subnetworks, many of which were circle-type subnetworks, consisting of one or two circle-like structures. Based on the number of nodes, edges, and the average number of neighbors (Supplementary Fig. 2), the subnetworks can be divided into three scales. Subnetwork 1 had the largest scale, including 15.5% of nodes and 38% of edges, with a high average number of neighbors (81). Subnetworks 2–11 were medium-scale subnetworks, including more than half of the nodes and edges. Finally, subnetworks 12–28 were small-scale subnetworks, with much fewer nodes and edges. Supplementary Table 2 shows the average numbers of topological parameters of the subnetworks. The scales of subnetworks were positively correlated with the average clustering coefficient and were negatively correlated with the network heterogeneity. Another interesting result is that the subnetworks exhibited different characteristics from the whole acidogenic network. The acidogenic network was relatively discrete, while subnetworks were more concentrated. For example, the network densities of the subnetworks were much higher than the whole acidogenic network (Supplementary Table 2), which means that subnetworks were more populated with edges than the whole acidogenic network.

Supplementary Fig. 3 and Supplementary Table 3 show the distribution of core features in different subnetworks. Although subnetwork 1 had the highest number of nodes, only 3 core features were located in it. 15 core features were distributed in medium-sized subnetworks and the other 2 core features were located in small subnetworks. Specifically, the medium-sized subnetworks 5 and 8 had the largest number of core features (4), followed by the large subnetwork 1 (3). Subnetwork 5 mainly contained butyrate (*Oscillibacter*, the ASV belonged to the phylum *Oscillospiraceae*, and *Caproiciproducens*) and propionate (*Paludibacter*) oxidizers[19,20], while subnetwork 8 included bacteria able to produce acetate, butyrate, and propionate (*Petrimonas*, *Bacteroides*, the ASV belonging to the family *Lachnospiraceae*, and NK4A214 group). On the other hand, core features in subnetwork 1 include fermenters and syntrophic bacteria that are able to produce acetate (*Mesotoga*, the ASV belonging to the order *Syntrophales*, and the ASV belonging to the family *Spirochaetaceae*). The core features that co-occurred in the same subnetwork might have similar metabolic preferences, which however should be confirmed by more solid evidence. In addition, *Clostridium sensu stricto* 12 and the ASV belonging to the phylum *Clostridiaceae* did not have a strong (Pearson correlation coefficient >0.7) correlation with the other core features and were located in two small subnetworks.

Table 1 summarizes the topological properties of the core features. Core features presented obvious variations in some of these topological properties, reflecting that they might have different interactions with other microbes and might play different roles in the microbial network. This is in line with the fact that core features were distributed within different subnetworks rather than clustered together. For example, the ASV belonging to the family *Spirochaetaceae* had 103 neighbors, connecting most of the genera in subnetwork 1, therefore was likely to be a hub in the subnetwork. In contrast, *Clostridium sensu stricto* 12 and *Christensenellaceae* R-7 group had only 3 and 2 neighbors, respectively. Thus, they did not interact with most of the other nodes in the network[23].

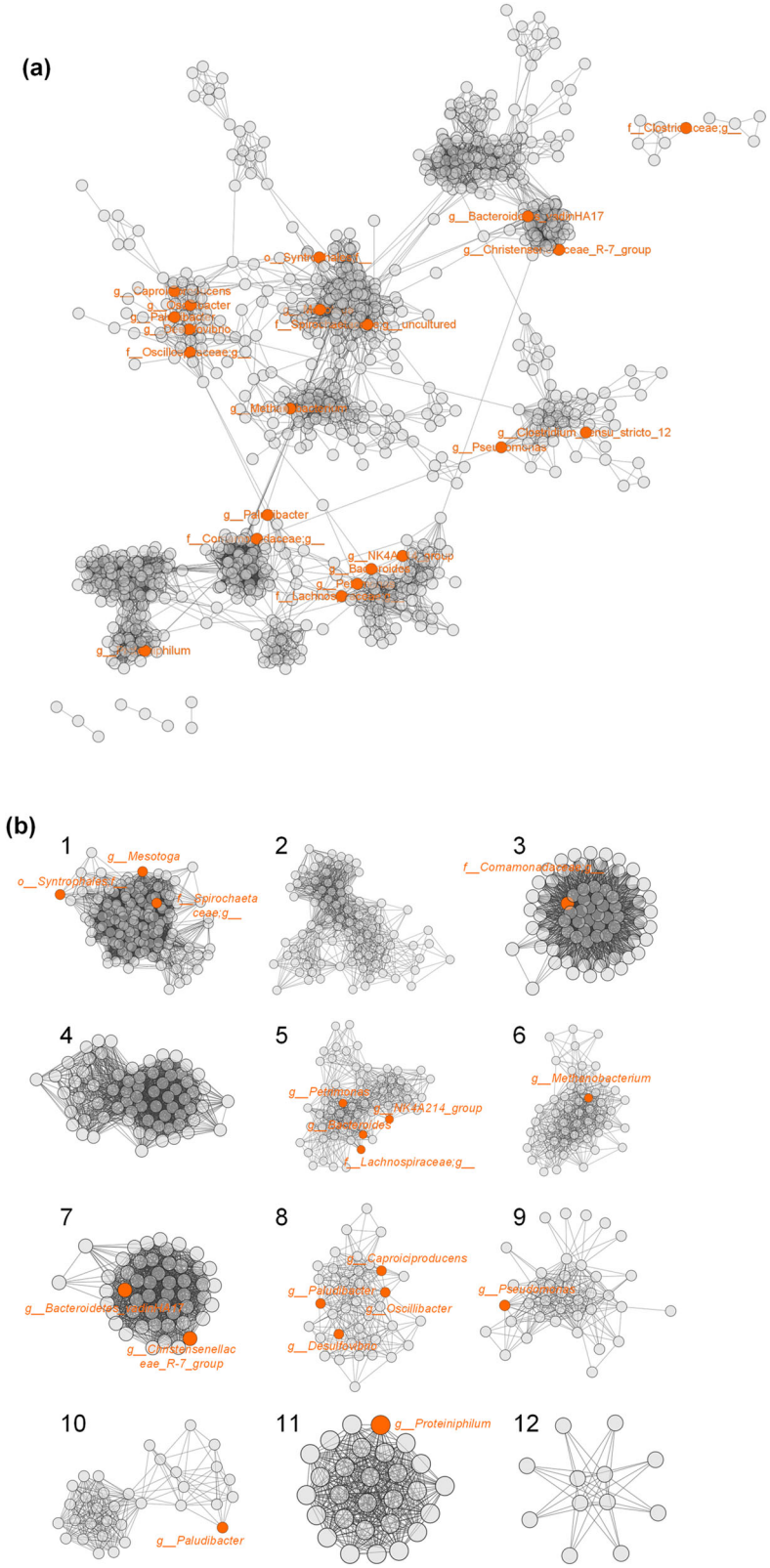


**Fig. 2 Acidogenic network and its subnetworks. a** Acidogenic network at the genus level (Pearson coefficient >0.7, $p \leq 0.05$) and **b** Top 12 subnetworks of the acidogenic network. The orange nodes represent core features.

Previous studies reported that inoculum samples had higher diversity than acclimated samples, but none of the studies had clarified this through the aspect of the microbial network. To assess this, the inoculum network was also established and was compared with the acidogenic network (Supplementary Fig. 4 and Supplementary Table 1). Generally, the inoculum network presented a distinct structure. However, the difference between the acidogenic network and the inoculum network was not in

**Table 1.** Topological properties of core features in the microbial network at the genus level.

| Core feature | Average shortest path length | Betweenness | Closeness | Clustering coefficient | Degree (no. of neighbor) | Topological coefficient |
|---|---|---|---|---|---|---|
| f__*Spirochaetaceae*;g__uncultured | 4.36 | 0.00 | 0.23 | 0.87 | 103.00 | 0.57 |
| g__*Paludibacter* | 6.04 | 0.00 | 0.17 | 0.74 | 13.00 | 0.44 |
| g__*Desulfovibrio* | 6.02 | 0.00 | 0.17 | 0.76 | 7.00 | 0.52 |
| f__*Oscillospiraceae*;g__ | 5.95 | 0.02 | 0.17 | 0.29 | 7.00 | 0.37 |
| g__*Oscillibacter* | 5.86 | 0.00 | 0.17 | 0.57 | 17.00 | 0.39 |
| g__*Caproiciproducens* | 6.09 | 0.00 | 0.16 | 0.81 | 13.00 | 0.47 |
| f__*Comamonadaceae*;g__ | 4.68 | 0.08 | 0.21 | 0.58 | 66.00 | 0.46 |
| g__*Macellibacteroides* | 6.04 | 0.01 | 0.17 | 0.76 | 7.00 | 0.39 |
| g__*Proteiniphilum* | 5.47 | 0.01 | 0.18 | 0.89 | 24.00 | 0.28 |
| g__*Mesotoga* | 4.56 | 0.01 | 0.22 | 0.82 | 65.00 | 0.49 |
| g__*Pseudomonas* | 5.47 | 0.08 | 0.18 | 0.73 | 12.00 | 0.25 |
| g__NK4A214 group | 6.28 | 0.00 | 0.16 | 0.65 | 21.00 | 0.39 |
| g__*Petrimonas* | 6.26 | 0.01 | 0.16 | 0.74 | 26.00 | 0.43 |
| g__*Bacteroidetes* vadinHA17 | 5.37 | 0.01 | 0.19 | 0.72 | 50.00 | 0.49 |
| g__*Christensenellaceae* R-7 group | 6.16 | 0.00 | 0.16 | 0.83 | 34.00 | 0.83 |
| g__*Clostridium sensu stricto* 12 | 10.45 | 0.00 | 0.10 | 1.00 | 3.00 | 0.52 |
| g__*Bacteroides* | 6.04 | 0.00 | 0.17 | 0.69 | 24.00 | 0.41 |
| g__*Methanobacterium* | 5.14 | 0.01 | 0.19 | 0.66 | 31.00 | 0.39 |
| o__*Syntrophales*;f__ | 5.08 | 0.00 | 0.20 | 0.67 | 12.00 | 0.39 |
| f__*Lachnospiraceae*;g__ | 6.26 | 0.03 | 0.16 | 0.85 | 15.00 | 0.38 |
| f__*Clostridiaceae*;g__ | 1.60 | 0.00 | 0.63 | 1.00 | 2.00 | 1.00 |
| Average of core features | 5.68 | 0.01 | 0.20 | 0.74 | 26.29 | 0.47 |
| Average of all genera | 5.72 | 0.01 | 0.19 | 0.70 | 32.79 | 0.51 |

the number of nodes (i.e., number of genera), but in the interacting relationships. Specifically, the inoculum network had 109% more edges and the average number of neighbors was also 132% higher than that of the acidogenic network. Therefore, genera in the inoculum samples had much more complex interactions than the acidogenic samples. The higher clustering coefficient and network density of the inoculum network also supported this result (Supplementary Table 1). These results indicate that acidogenic acclimation processes decreased the microbial interactions. One possible explanation is that compared with the inoculum samples, acidogenic samples have fewer microbial functions which were interlinked with interactions such as syntrophic oxidation of substrates.

## Rare community

In contrast with the dominant species, rare species greatly contribute to the richness but account for a low percentage of abundance in the microbial ecosystem[4]. To understand whether there was a rare community in the acidogenic systems investigated, rare features were defined and identified. Surprisingly, only one ASV belonging to the phylum *Firmicutes* was found to be the rare feature when the criterion of ubiquitous features was defined as the occurrence frequency in more than 70% of all samples. Rare features in acidogenic systems are more likely to exist stochastically, unlike the existing pattern of the core microbial community.

Decreasing the criterion of the occurrence frequency from 70 to 50% led to the discovery of 5 rare features, including *Candidatus* Caldatribacterium, *Phaselicystis*, *Thermovirga*, one ASV belonging to the family *Hungateiclostridiaceae*, and one belonging to *Firmicutes* (Fig. 3). *Candidatus* Caldatribacterium, *Phaselicystis*, *Thermovirga*, and family *Hungateiclostridiaceae* are fermenting bacteria, able to function in VFAs production[24–27].

Rare species play different roles in ecosystems. Generally, rare species serve as a seed bank with huge ecological potential which can respond to environmental changes and maintain system stability[28]. Besides, rare species are responsible for functions such as the removal of micropollutants[29]. Although the relative abundance of rare features was low, the fact that 4 of 5 rare features were able to function in VFAs production implied a possibility that they might to some extent contribute to acidogenesis which however needs further investigation.

## Top metabolic pathways

PICRUSt2 was applied to predict metabolic pathways of acidogenic systems based on 16 S rRNA marker genes. The NSTI values of acidogenic samples ranged from 0.03 to 0.20 with a mean value of 0.12. Low NSTI values suggested that these samples can be tracked for PICRUSt2 prediction[30]. A total of 410 pathways from the MetaCyc database were predicted in all acidogenic samples. Figure 4a shows the top 40 pathways, including 33 biosynthesis pathways, 2 pathways associated with degradation/utilization/assimilation, and 5 pathways associated with the generation of precursor metabolites and energy. Biosynthesis pathways constitute the cells' spectrum of biosynthetic capacities, including the pathways of synthesis of small molecules, macromolecules, and cell structure components. Most (85%) of the top 33 biosynthesis pathways belonged to amino acid biosynthesis, fatty acid and lipid biosynthesis, and nucleoside and nucleotide biosynthesis. Similar pathways were reported to be dominant in fermentation systems[5,31]. It should be noted that only 2 fermentation pathways were found in the top 40 pathways, including one pathway associated with "pyruvate fermentation to acetate and lactate II" and another one associated with "pyruvate fermentation to isobutanol".

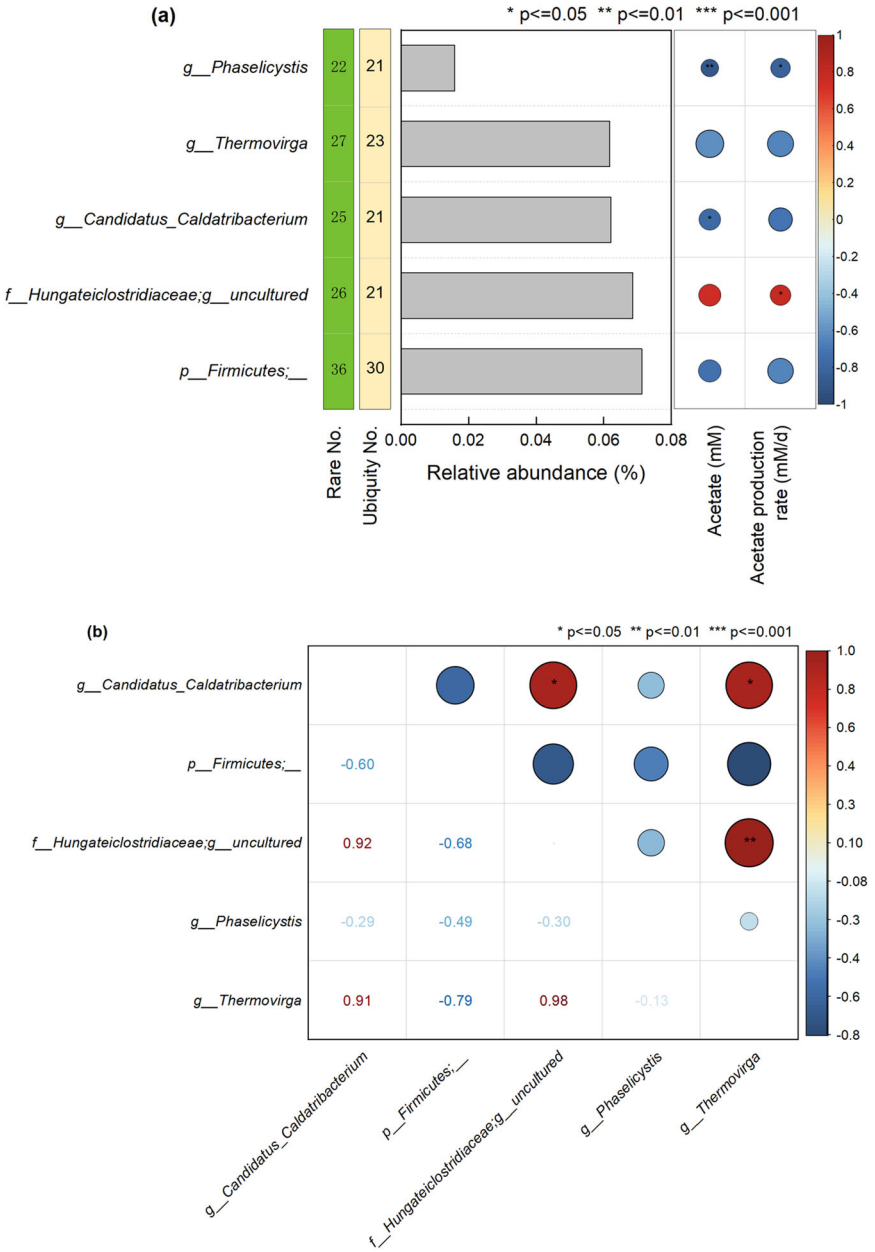


**Fig. 3 Characterizations of rare features. a** The taxonomic composition of rare features at the genus level and their frequency of abundance (green column), frequency of ubiquity (yellow column), and correlation with acetogenic performance and **b** Pearson correlation among rare features.

## Acid-producing pathways

Apart from the 2 fermentation pathways mentioned above, another 13 pathways associated with acidogenesis were further identified in the acidogenic samples (Fig. 4b), which in total accounted for 3.6% of the total number of all metabolic pathways identified. The acidogenic samples were clustered into two groups according to the abundance of these 15 pathways (Fig. 4b), with one group containing the high abundance of pathways associated with "pyruvate fermentation to isobutanol (engineered)", "pyruvate fermentation to acetate", "pyruvate fermentation to acetate and lactate II", "homolactic fermentation", and "acetyl-CoA fermentation to butanoate II", and another group only containing high abundance of pathways associated with "pyruvate fermentation to isobutanol (engineered)". The acidogenic samples in these two groups, however, could not be clearly distinguished by the category of substrates or products. Pyruvate, which was correlated with 5 of 15 pathways, seemed to be a key intermediate in the acidogenic process. According to the MetaCyc database, all $CO_2$, lactate, and components of brewery spent grain (e.g., glucose) could be converted to pyruvate via multiple reactions. Therefore, one strategy to promote the targeted acid or alcohol yield is to increase the pyruvate pool by knocking out genes that contributed to by-product formation using pyruvate as the precursor. For example, to increase the isobutanol production from pyruvate in engineered *Escherichia coli*, the genes that regulated the by-product formation from pyruvate were deleted and levels of pyruvate available for isobutanol production were increased[32].

As shown in Supplementary Fig. 5, four $CO_2$ fixation pathways were identified in the acidogenic samples investigated. Supplementary Table 4 summarizes the reactants, products, and subsequent pathways of $CO_2$ fixation. The "Calvin Benson

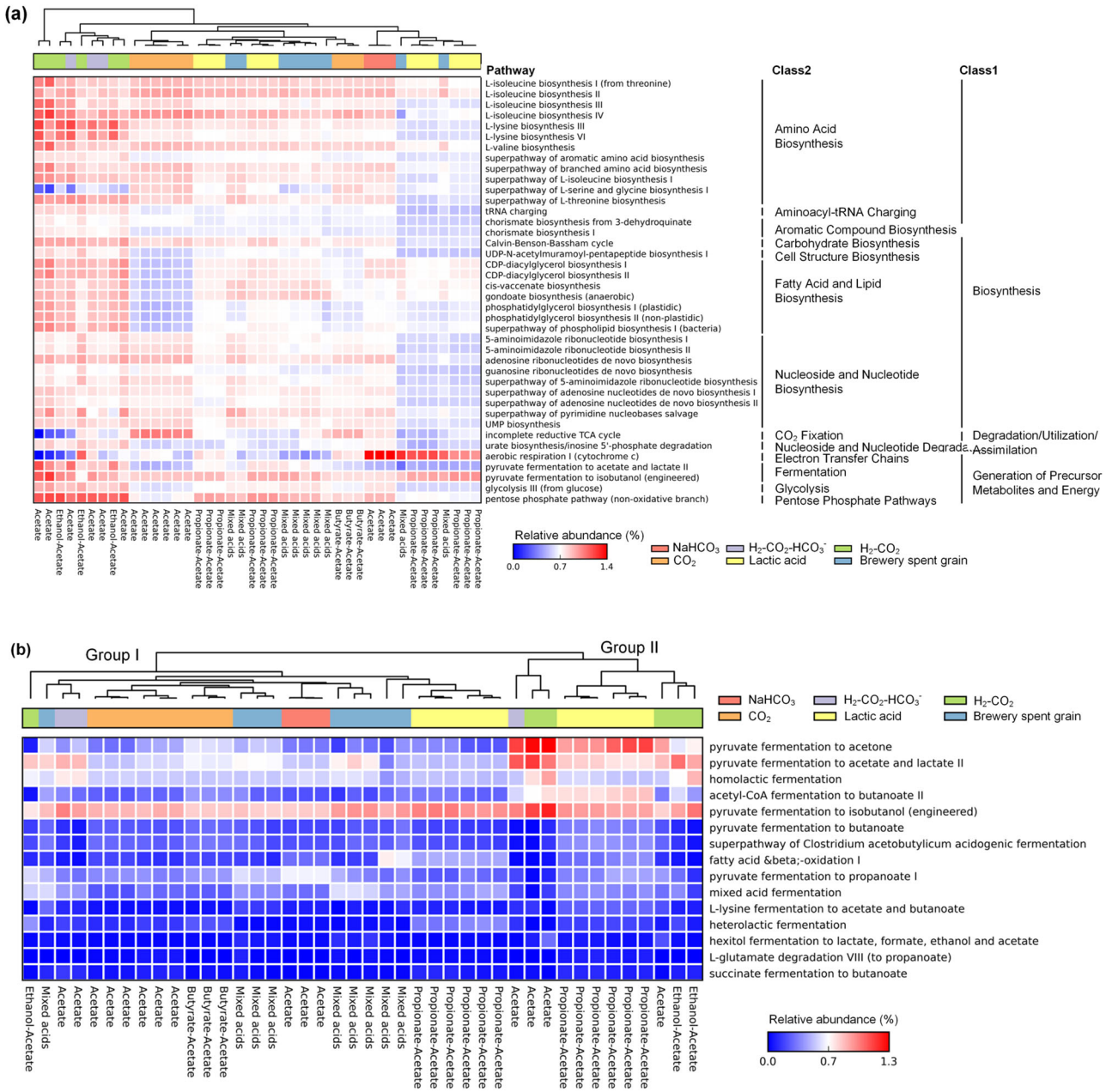


**Fig. 4 Top metabolic pathways and acid-producing pathways. a** Relative abundance of top 40 predicted pathways and **b** relative abundance of fermentation pathways.

Bassham cycle" turned out to be the most abundant $CO_2$ fixation pathway, being responsible for generating 3-phospho-D-glycerate, glycerone phosphate, or D-glyceraldehyde 3-phosphate which could be used in the biosynthetic metabolism of autotrophic bacteria. The "reductive tricarboxylic acid (TCA) cycles I and II" from autotrophic eubacteria and archaea were also found in acidogenic samples. The "reductive TCA cycles I and II" are commonly found in autotrophic eubacteria and archaea. The basic cycle results in the fixation of 2 molecules of $CO_2$ and the generation of 1 molecule of acetyl-CoA[33]. Another molecule of $CO_2$ can be fixed by the carboxylation of acetyl-CoA to produce pyruvate[34]. The "reductive acetyl coenzyme A pathway", known as the Wood-Ljungdahl pathway, is the most effective nonphotosynthetic $CO_2$ fixation pathway by acetogens[35,36]. Two molecules of $CO_2$ can be converted to acetyl-CoA in the Wood-Ljungdahl pathway. Acetate, ethanol, and butyrate could be produced from $CO_2$ through the detected Wood-Ljungdahl pathway coupled with the acetyl-CoA reduction in acidogenic samples[9,16,20].

## Roles of core features in acidogenic metabolism

To examine the contribution of core features to metabolic pathways, genes involved in the top 40 pathways as well as the 15 fermentation pathways were further assigned to ASVs. As shown in Supplementary Fig. 6a, 4.6–74.3% of the top 40 pathways belonged to core features, with an average of 26.2% per pathway. The highest average percentage of core feature assignment was found in samples producing ethanol and acetate, accounting for an average of 45.2% per pathway, while the lowest one was found in samples producing butyrate and acetate (20%). Similar results were found in fermentation pathways where core

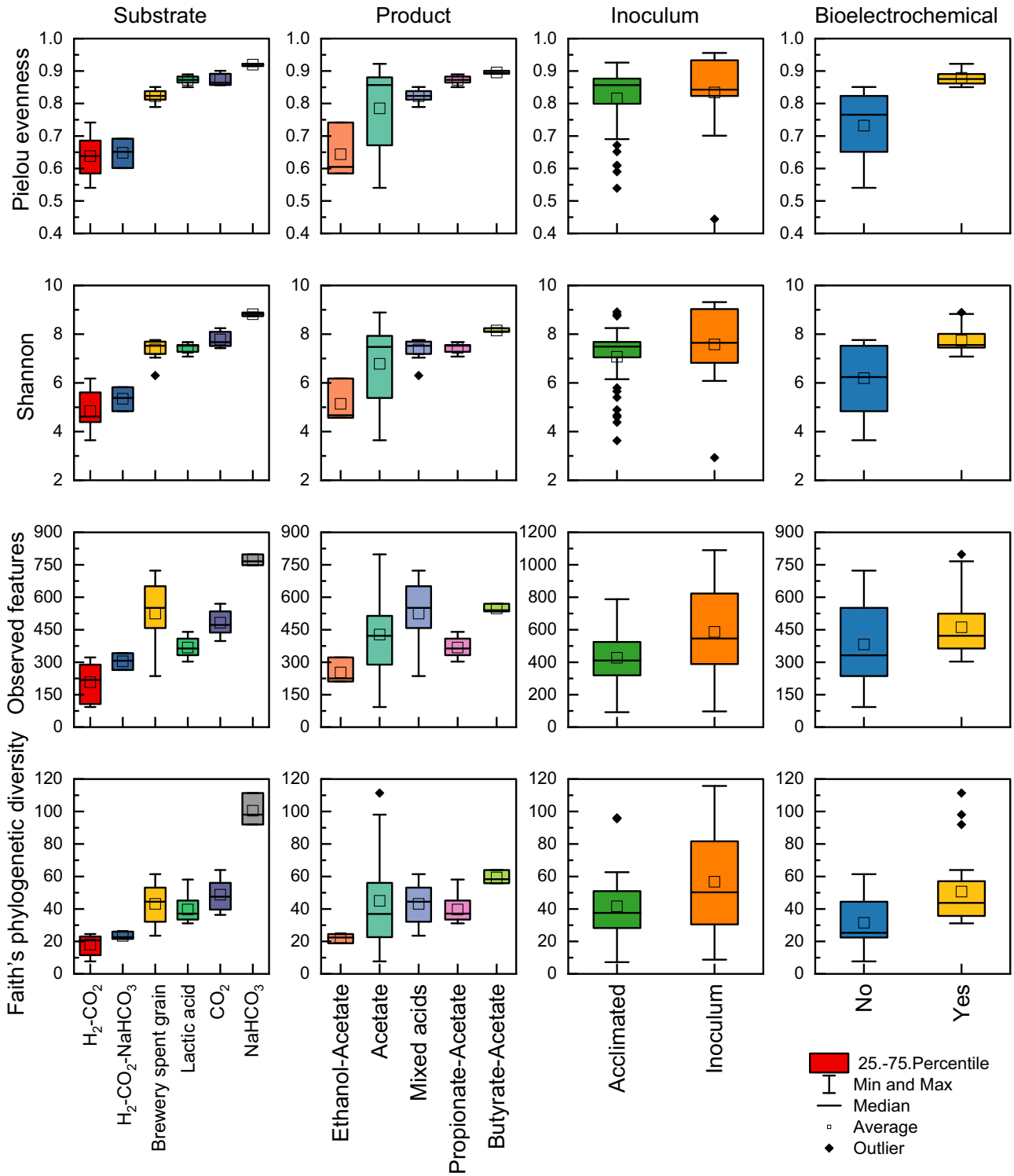


**Fig. 5 Alpha diversity indexes (Pielou evenness, Shannon, observed features, and Faith's phylogenetic diversity) of acidogenic samples grouped by substrate, product, inoculum, and bioelectrochemical characteristic.** The line and square in the box mark the mid-point and average values of the data, respectively. The upper and lower edges of the box mark that the seventy-five and twenty-five of the data fall below the upper quartile and the lower quartile, respectively. The upper and lower whiskers represent the maximum and minimum values respectively, excluding outliers. The black diamond represents the outlier.

features accounted for 4.9–66.1% of the fermentation-pathway genes (Supplementary Fig. 6b), with an average of 27.0% per pathway. This result implies that the core features participated not only in the top 40 pathways, but also in the 15 fermentation pathways in the acidogenic samples.

### Microbial diversity

Alpha diversity indexes associated with evenness (Pielou's evenness and Shannon's index) and richness (Observed ASVs and Faith's phylogenetic diversity) were calculated to investigate the microbial diversity of the acidogenic samples (Fig. 5).

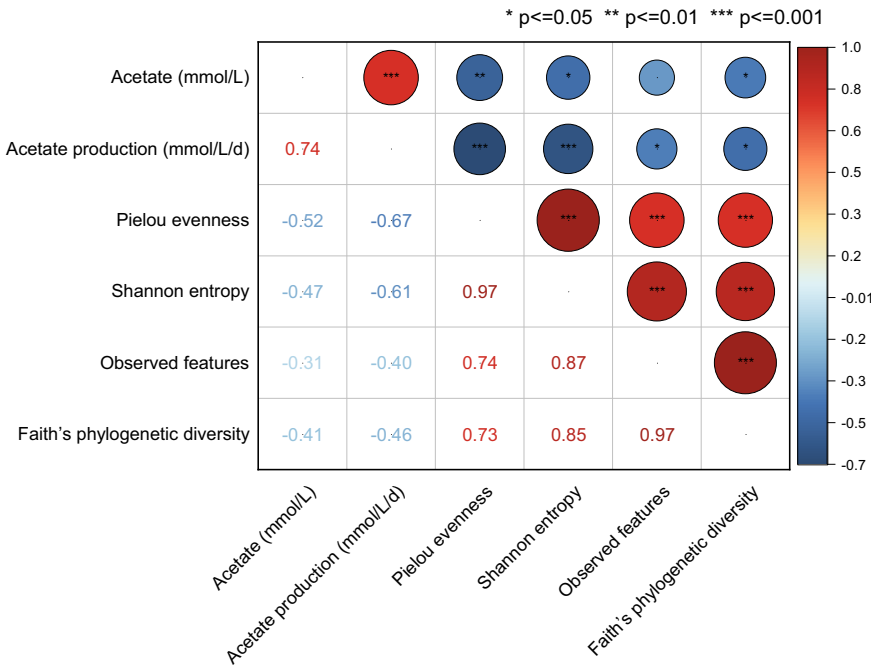


**Fig. 6 Pearson correlation between alpha diversity indexes and acetogenic performances.** The size of the circle represents the size of the absolute value of the index and the color of the circle represents the value of the index. Red is positive and blue is negative. The star represent the significant value.

When grouped by substrates, samples fed with $NaHCO_3$ showed the highest level of both evenness and richness, while samples fed with $H_2$–$CO_2$ demonstrated the lowest diversity. Particularly, diversity of samples fed with $NaHCO_3$ was significantly higher than that of samples fed with $H_2$–$CO_2$, $CO_2$, and $H_2$–$CO_2$–$NaHCO_3$ ($p \le 0.05$) (Supplementary Table 5). On the other hand, when grouped by products, sludge samples producing butyrate and acetate showed the highest level of diversity, while those with ethanol and acetate as main products presented the lowest diversity (Supplementary Table 6). Compared with acidogenic samples, the diversity indexes of inoculum samples were relatively higher, of which the observed ASVs index was significantly different ($p \le 0.01$). This might be due to the washout or decay of some microorganisms that were not tolerant to the low pH in the acidogenic systems or the immigrated microorganisms from the substrates lead to the high microbial diversity of the inoculum samples[3].

It is generally accepted that a high diversity increases functional resilience under stress conditions and results in better system performance[37,38]. However, as shown in Fig. 6, all alpha diversity indexes were negatively correlated with the acetogenic performance. This may be because that acetogenesis is less dependent on the diversity. For example, when $CO_2$ and $H_2$ are fed, homoacetogenesis occurs spontaneously without interacting with other microbes, thus increasing the number of other microorganisms involved. Tao et al.[3] also reported that only a weak correlation between diversity indices and performance parameters was found in anaerobic digestion samples.

To understand community assembly patterns, PCoA was applied to investigate the variations in community composition (beta diversity) based on taxonomic (Jaccard distance and Bray–Curtis distance) and phylogenetic (unweighted and weighted UniFrac distance) diversity (Fig. 7). Considerable variations between acidogenic samples were observed at both the taxonomic and phylogenetic levels. Among them, the beta diversity results based on the weighted Unifrac distance explained the highest variation (38.36%). Microbial communities of acidogenic samples grouped by substrate or products exhibited significant differences ($p = 0.001$) (Fig. 7), suggesting that these two parameters also affected the microbial community composition.

### Effects of environmental variables on the core microbial community

As demonstrated by CCA (Fig. 8a), environmental variables (e.g., electrode potential, substrate, and pH) and acetogenic performances had a clear impact on the core features. Together, environmental variables and acetogenic performance explained 73.5% of the core feature variation ($p \le 0.05$) (Supplementary Table 7). Surprisingly, the electrode potential was the key contributing factor ($p \le 0.01$), accounting for 48.2% of the explanation. Core features responded to these factors variously. Many core features were located near the arrow of the bioelectrochemical factor. This may be because extra electrons provided by the electrodes can motivate energetically unfavorable reactions[11], and may thus favor the growth of some acidogenic bacteria. Specifically, the distribution of 16 core features was significantly different ($p \le 0.005$) in samples with or without the electrode potential (Supplementary Table 8). For example, *Clostridium sensu stricto* 12, *Bacteroides*, *Bacteroidetes vadinHA17*, and the ASV belonging to the family *Lachnospiraceae* were more abundant in samples without the electrode potential (Supplementary Table 8). In contrast, *Methanobacterium*, *Desulfovibrio*, *Oscillibacter*, and *Caproiciproducens* were dominant in the samples with the electrode potential (Supplementary Table 8). *Methanobacterium* species could produce $H_2$ instead of $CH_4$ when confronted with an excess supply of reducing equivalents on a cathode in an acetate-producing bioelectrochemical system[39]. Furthermore, *Desulfovibrio* and *Pseudomonas* (acetogen) are commonly recognized as electroactive bacteria, able to accept electrons from cathodes in microbial electrolysis cells[40,41]. In addition, from the subnetwork aspect (Fig. 2 and Supplementary Table 3), all core features located in subnetwork 8 were favored by the electrode

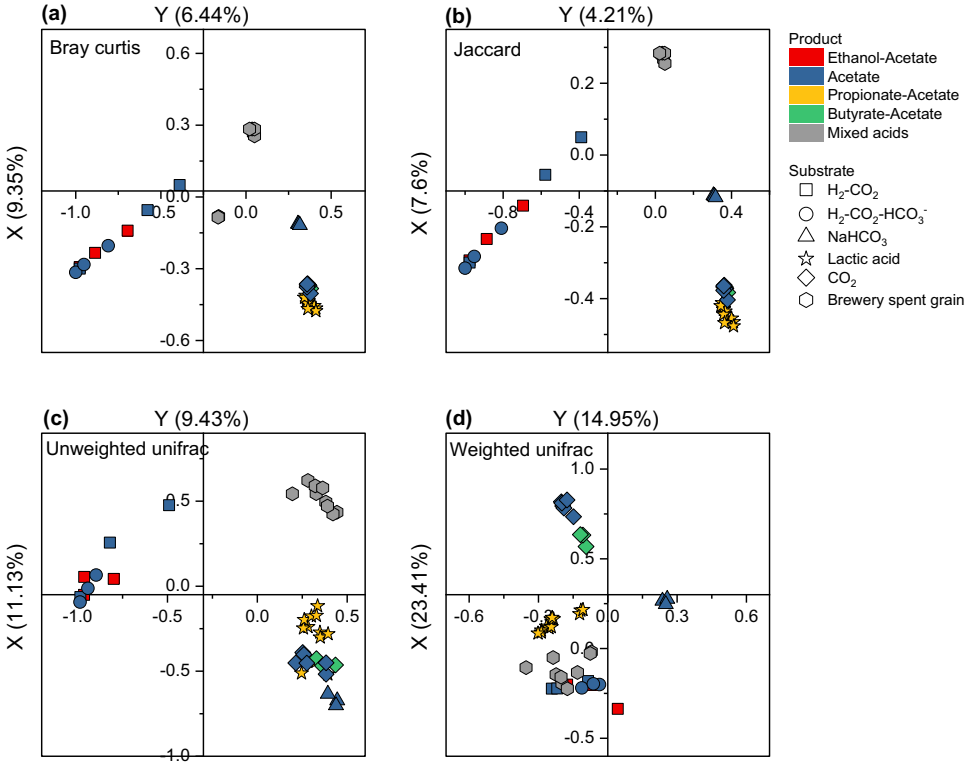


**Fig. 7 PCoA of beta diversity. a** Bray–Curtis distance, **b** Jaccard distance, **c** unweighted UniFrac distance, and **d** weighted UniFrac distance grouped by product (color) and substrate (shape).

potential. Environmental variables may also play a role in the establishment of subnetworks.

On the other hand, *Clostridium sensu stricto* 12 and the ASV belonging to the phylum *Clostridiaceae* were located far away from other core features and near the arrow of the acetate-production rate (Fig. 8a), confirming their dissimilarity of distribution and their important roles in maintaining acetogenic performance. Besides, although *Methanobacterium* was identified as one of the core features in acidogenic samples, the CCA result shows that it was not closely correlated with most of the other core features and had a distinct pH preference.

### Effects of environmental variables on the metabolic pathways

The RDA further revealed that the top 40 metabolic pathways, 15 fermentation pathways, and 4 $CO_2$ fixation pathways responded differently to the environmental variables and gave a different acetogenic performance (Fig. 8b) which explained a total of 65.3% of the pathway variation ($p \le 0.05$, contribution of acetate concentration excluded). The bioelectrochemical factor was the highest contributing variable ($p \le 0.01$), accounting for 49.3% of the explanation, followed by pH (25.2%) (Supplementary Table 9). Furthermore, more than half of the fermentation pathways were positively correlated with the bioelectrochemical factor even though the bioelectrochemical factor was negatively correlated with acetogenic performance, implying that the electrode potential might benefit the production of other VFAs rather than acetate, which needs further confirmation. For example, propionate generation from pyruvate (P108-PWY), which can be first produced during lactate fermentation[42], was positively correlated with the bioelectrochemical factor. Another intriguing result is that most of the top 40 pathways were clustered together and had a high positive correlation with the acetate-production rate, but were negatively correlated with the bioelectrochemical factor. Since 33 of the top 40 pathways are related with biosynthesis, it is also necessary to confirm in the future studies whether the electrode potential application negatively impacts the growth of microbes.

### Effects of environmental variables on the diversity

Alpha diversity indexes of acidogenic samples were distinguished by their pH ($p \le 0.01$). As shown in Supplementary Fig. 7, samples with a low pH tended to harbor low microbial diversity and vice versa. Some samples with pH lower than 4.5 were those producing ethanol via solventogenesis. This is in line with the fact that solventogenesis usually occurs at a pH ranging from 4.5 to 5.5[9]. Furthermore, these samples were fed with $H_2$–$CO_2$ but not $NaHCO_3$, consistent with the result from Dessi et al.[16] that changing carbon source from bicarbonate to $CO_2$ reduced the alpha diversity of acidogenic systems. This may be also because bicarbonate was able to act as a buffer, avoiding the pH drop and eventually retaining more microbes that are sensitive to pH. On the other hand, the temperature which is also a key factor affecting the microbial population[8] did not show a significant effect on diversity ($p > 0.05$) (Supplementary Fig. 8). Acidogenic samples cultivated at 25 and 37 °C might share relatively similar diversity.

Another intriguing result is that the electrode potential exerted significant effects on the diversity ($p \le 0.05$). Sludge samples collected from the bioelectrochemical systems showed a higher diversity than those without the electrode potential. It is recognized that electrochemical control affects the extracellular and intracellular metabolism of fermentative microorganisms[10]. Figure 5 further shows that the electrode potential could benefit the diversity of acidogenic samples, which might be partially due to the enrichment of electroactive bacteria. Moreover, when considering beta diversity, samples classified as with or without electrode potential showed a clear distinction that could be observed ($p = 0.001$) (Supplementary Fig. 9). This may be because

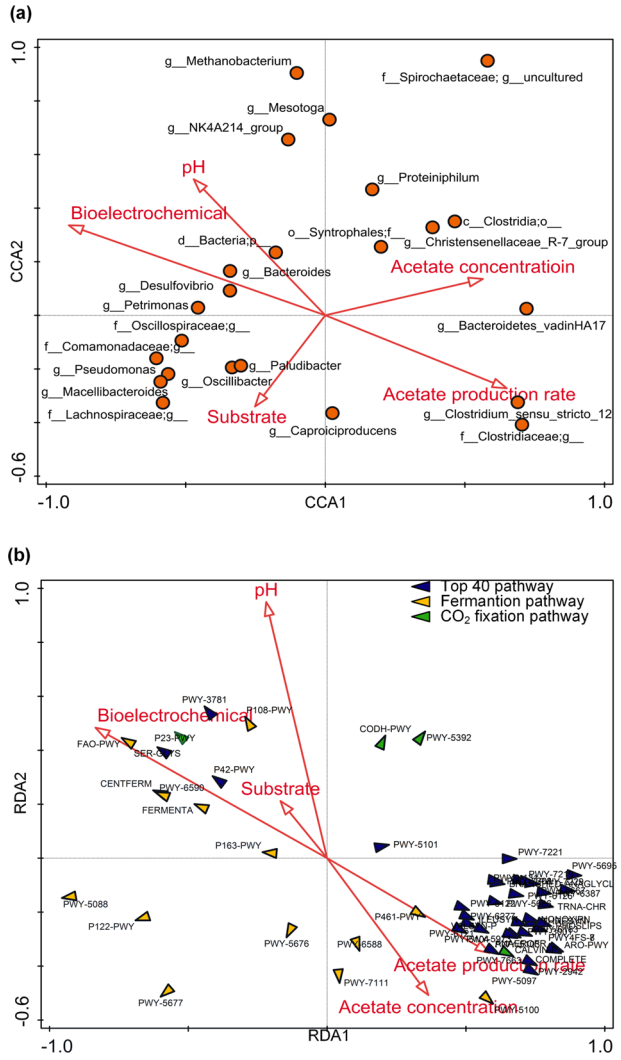


**Fig. 8 Effects of environmental variables on core features and predicted pathways. a** Canonical correlation analysis (CCA) of core features, environmental variables, and acetogenic performance as well as **b** redundancy analysis (RDA) of predicted pathways, environmental variables, and acetogenic performance. The full name of each pathway can be found on MetaCyc database (https://metacyc.org/).

the electrode potential on the cathodes can provide a unique microbial niche by establishing interactions either between microorganisms or between microorganisms and the electrode surface[10], both of which can result in a shift of the community composition.

## Implications

The above results provide insights into the optimization of acidogenic processes. A core microbial community was identified in the studied acidogenic samples. Since most of these core features are acidogenic bacteria and harbored a large percentage of predicted genes involved in acid-producing pathways, they are likely to be the main acid producers in many acidogenic systems for anaerobic biorefinery of valuable bioproducts. From the network aspect, the decentralized distribution of core features in different subnetworks suggested that many of them did not have co-occurring relationship and tended to play different roles in the acidogenic network. Although the exact function of each subnetwork remains unknown, the existence of many high-density subnetworks might imply a potential subnetwork-based regulation strategy of the acidogenic system.

Although diversity is suggested to play important roles in system performance[37,38], in this study, all alpha diversity indexes were negatively correlated with the acetogenic performance. This may be because acidogenic processes are usually operated on relatively low pH conditions which are not suitable for many microbes. Therefore, the importance of the high microbial diversity to acidogenic performance should be reevaluated. The distinct differences between the inoculum network and the acidogenic network also supported the hypothesis that microbial diversity plays a minor role in acidogenic systems. The acclimation did not obviously affect the number of microbes (diversity) that have a strong positive correlation but significantly decreased their interacting relationships. Instead, core features and their response to environmental changes should be focused on when optimizing the acidogenic performance.

The bioelectrochemical factor has the greatest influence on not only core features but also the fermentation pathways. More than half of the core features and fermentation pathways were positively correlated with the electrode potential. Therefore, when aiming to enrich the associated core features or produce the associated products, microbial electrosynthesis may be a promising approach. Since some of the core features that positively affected by the electrode potential were located in the same subnetwork, a hypothesis is that the bioelectrochemical factor is able to regulate the microbial community at the subnetwork level but this needs further verification. Nevertheless, it should also be noted that most of the top 40 metabolic pathways, and some other core features were negatively correlated with the bioelectrochemical factor. Thus, in future studies, the accurate effects of the bioelectrochemical factor on shaping the microbial community should be further examined as well.

This study investigated the core microbial community, microbial network, metabolic pathways, and microbial diversity of acidogenic samples. A total of 23 microbes were identified as the core features, with *Clostridium* as the most abundant one. The non-centralized distribution of core features in different subnetworks and their topological parameters suggested that they play different roles in the microbial network. On average, 27% of the predicted fermentation-pathway genes belonged to the core microbial community, implying their important roles in acidogenesis. Electrode potential contributed the most to the distribution of both the core microbial community (48.2%) and metabolic pathways (49.3%). More than half of the core features and fermentation pathways were positively correlated with the electrode potential. Posing an electrode potential, as e.g., in microbial electrosynthesis, may thus be a promising tool to shape an anaerobic digestion microbiome towards a VFAs and alcohols one.

# METHODS

## Sample description

A total of 60 biomass samples were collected from four different experiments with acidifying bioreactors conducted at National University of Ireland Galway, Ireland. The reason to choose samples from these experiments was that they were conducted under different environmental variables including different inoculum, temperature, pH, substrate, and electrode potential. The wide variety of environmental variables allowed us to have a more general profile regarding how they affect the microbial community. Detailed sample information is summarized in Table 2. Specifically, twelve samples were collected from group I[9]. The inoculum of group I was collected from an up-flow-anaerobic sludge bed (UASB) reactor producing methane from dairy wastewater. Acidogenic samples were collected from batch-test reactors operated at 25 °C or 37 °C, fed with $H_2$–$CO_2$, $H_2$, $CO_2$, or $NaHCO_3$ as the substrate. Samples of group I were

**Table 2.** Details of the 60 samples collected from the four experiments.

| Group | Inoculum | No. of samples | No. of temperature | No. of substrate type | Electrode potential | Experiment description | Reference |
|---|---|---|---|---|---|---|---|
| I | Up-flow-anaerobic sludge bed (UASB) reactor producing methane from dairy industry effluent | 3 inoculums/ 9 acidogenic samples | 6 25 °C/ 3 37 °C | 6 $H_2 + CO_2$/ 3 $H_2 + CO_2 + NaHCO_3$ | – | Batch tests were conducted in 125 mL serum bottles with 50 mL medium (gas: liquid ratio of 3:2) and granular sludge with an initial VS concentration of 1.0 g/L. | [9] |
| II | Anaerobic digester treating cheese processing wastewater | 3 inoculums/ 12 acidogenic samples | 12 25 °C | 12 Lactic acid | −1.0 V | Bioelectrochemical reactor operated for 225 days at an applied cathode potential of −1.0 V vs. Ag/AgCl. | [20] |
| III | Anaerobic digester treating cheese processing wastewater | 3 inoculums/ 12 acidogenic samples | 12 25 °C | 9 $CO_2$/3 $NaHCO_3$ | −1.0 V | Bioelectrochemical reactor operated for 25 days at an applied potential of −1 V vs. Ag/AgCl. | [16] |
| IV | 3 Rumen fluid/3 Rhino manure/3 Tiger manure | 9 inoculums/ 9 acidogenic samples | 9 37 °C | 9 Brewery spent grain | – | Batch tests were conducted in serum bottles with VS concentration of 6.0 g/L. | [43] |

collected at the end of the batch-test incubation. The main products of these samples were ethanol and/or acetate. Fifteen samples were collected from group II[20]. The inoculum of group II was collected from an anaerobic digester treating cheese processing wastewater (Dairygold, Cork, Ireland). Acidogenic samples were collected from microbial electrosynthesis reactors producing propionate and acetate from lactic acid at 25°C and an applied cathodic potential of −1.0 V vs. Ag/AgCl. Samples of group II were collected at the end of the electrosynthesis experiment. Another fifteen samples were collected from group III[16]. The inoculum was digested sludge from a dairy processing industry (Dairygold, Ireland). Acidogenic samples were collected from microbial electrosynthesis reactors fed with $CO_2$ or $NaHCO_3$ at 25 °C and the applied cathodic potential of −1.0 V vs. Ag/AgCl. Samples of group III were collected on days 47, 183, 206, and 225 of reactor operation (total operation of 225 days). The main products of these samples were acetate and/or butyrate. Eighteen samples were collected from group IV[43]. The inoculum samples included three from rumen fluid, three from rhino manure, and three from tiger manure. Acidogenic samples were collected on day 21 from batch-test reactors operated at 37°C degrading brewery spent grain (total operation of 30 days)[43].

DNA was extracted using DNA extraction kits following the manufacturer's protocol[9,16,20,43]. 16S rRNA genes were amplified using the universal primer pair 515F/806R. Library preparation and high throughput sequencing were conducted on the Illumina Miseq platform.

## Microbial community analysis

Amplicon sequence variants (ASVs) were constructed using the Qiime2 workflow[44]. Briefly, sequence merging, quality filtering, ASVs table construction, and taxonomic assignment were performed[44]. Sequence merging was conducted using PEAR with a minimum overlap of 10 bps and a minimum assembled length of 200 bps. Data2 was applied for sequence quality filtering and ASVs table construction[45]. Taxonomic analysis was performed by comparing with the SILVA database (silva-138-99-nb-classifier). The definition of the core microbial community was based on Wu et al.[2] with a slight modification. The criterion of "overall abundant features" was defined as ASVs with a total relative abundance higher than 0.1% of all acidogenic samples. The criterion of "ubiquitous features" was defined as ASVs with an occurrence frequency in more than 70% of all acidogenic samples. The criterion of "frequently abundant features" was defined as ASVs abundant in at least 50% of the acidogenic samples. ASVs fulfilling all the above three criteria were defined as core features in acidogenic systems. The rare community was defined similarly except that the criteria of "overall rare features" was defined as ASVs with the total relative abundance lower than 0.1% and the criteria of "frequently rare features" was defined as ASVs rare in at least 50% of the samples.

## Microbial metabolic pathways prediction using PICRUSt2

PICRUSt2 was used to predict the metabolic pathways of the microbial communities based on 16S rRNA marker gene profiles[46]. The functional annotation of PICRUSt2 predictions was obtained based on the MetaCyc database[47]. The nearest sequenced taxon index (NSTI) was calculated to evaluate the reference genome coverage for accurate PICRUSt prediction. Estimated MetaCyc pathways related to the functional potential involved in acidogenesis were manually classified, based on their MetaCyc identification. Predicted pathways were assigned to ASVs for the clarification of microbes involved in metabolic pathways.

## Statistical analysis

Alpha and beta diversities were calculated by Qiime2 diversity plugin using the extracted ASVs[44]. The Kruskal–Wallis test was used to examine the significance of the alpha diversity[48]. The significance of the beta diversity was evaluated by the PERMANOVA test (pairwise comparison and 999 permuations). Principle coordinates analysis (PCoA) was applied to investigate the beta diversity. Heatmaps of metabolic pathways were generated using the STAMP software[49]. Detrended correspondence analysis (DCA), canonical correlation analysis (CCA), and redundant analysis (RDA) were conducted by the Canoco5 software. To normalize the types of substrates for statistical analysis, carbon substrates were calculated based on the molecular mass (brewery spent grain excluded). Since the length of axis 1 of DCA is lower than 2, RDA was selected to evaluate the correlation between the metabolic pathways and environmental variables[50]. Pearson correlation analysis was performed using the Origin software. Network analysis was performed in Cytoscape and Origin software based on the Pearson correlation coefficients (Pearson correlation coefficients >0.7, $p$ value ≤ 0.05). Subnetworks were extracted using the CytoCluster plugin[51]. A $p$ value ≤ 0.05 was considered statistically significant.

## DATA AVAILABILITY

All treated data that support the findings of this study are included in the present article. Raw datasets can be obtained from the corresponding author upon request. The sequences generated were deposited in the NCBI Sequence Read Archive (SRA) with accession number PRJNA853100.

Received: 22 November 2021; Accepted: 15 July 2022;
Published online: 06 September 2022

## REFERENCES

1. Strazzera, G., Battista, F., Garcia, N. H., Frison, N. & Bolzonella, D. Volatile fatty acids production from food wastes for biorefinery platforms: a review. *J. Environ. Manag.* **226**, 278–288 (2018).
2. Wu, L. et al. Global diversity and biogeography of bacterial communities in wastewater treatment plants. *Proc. Natl Acad. Sci. U.S.A.* **4**, 1183–1195 (2019).
3. Tao, Y. et al. Biogas productivity of anaerobic digestion process is governed by a core bacterial microbiota. *Chem. Eng. J.* **380**, 122425 (2020).
4. Yin, Q., Wang, Z. & Wu, G. Impacts of environmental factors on microbial diversity, distribution patterns and syntrophic correlation in anaerobic processes. *Arch. Microbiol.* **201**, 603–614 (2019).
5. Zhang, Q. et al. Shifts of microbial community and metabolic function during food wastes and waste activated sludge co-fermentation in semi-continuous-flow reactors: Effects of fermentation substrate and zero-valent iron. *Bioresour. Technol.* **313**, 123686 (2020).
6. Park, J. H. et al. Metabolic flux and functional potential of microbial community in an acidogenic dynamic membrane bioreactor. *Bioresour. Technol.* **305**, 123060 (2020).
7. Liu, H. et al. Acidogenic fermentation of proteinaceous sewage sludge: Effect of pH. *Water Res.* **46**, 799–807 (2012).
8. Liu, C. et al. The effects of pH and temperature on the acetate production and microbial community compositions by syngas fermentation. *Fuel* **224**, 537–554 (2018).
9. He, Y., Cassarini, C., Marciano, F. & Lens, P. N. L. Homoacetogenesis and solventogenesis from $H_2/CO_2$ by granular sludge at 25, 37 and 55 °C. *Chemosphere* **265**, 128649 (2021).
10. Moscoviz, R., Toledo-Alarcon, J., Trably, E. & Bernet, N. Electro-fermentation: how to drive fermentation using electrochemical systems. *Trends Biotechnol.* **34**, 856–865 (2016).
11. Moscoviz, R., Trably, E. & Bernet, N. Electro-fermentation triggering population selection in mixed-culture glycerol fermentation. *Microb. Biotechnol.* **11**, 74–83 (2018).
12. Vassilev, I. et al. Microbial electrosynthesis of isobutyric, butyric, caproic acids, and corresponding alcohols from carbon dioxide. *ACS Sustain. Chem. Eng.* **6**, 8485–8493 (2018).
13. Baleeiro, F. C. F., Kleinsteuber, S. & Strauber, H. Hydrogen as a co-electron donor for chain elongation with complex communities. *Front. Bioeng. Biotechnol.* **9**, 650631 (2021).
14. Shen, Y., Brown, R. C. & Wen, Z. Syngas fermentation by *Clostridium carboxidivorans* P7 in a horizontal rotating packed bed biofilm reactor with enhanced ethanol production. *Appl. Energy* **187**, 585–594 (2017).
15. Sun, X., Atiyeh, H. K., Zhang, H., Tanner, R. S. & Huhnke, R. L. Enhanced ethanol production from syngas by *Clostridium ragsdalei* in continuous stirred tank reactor using medium with poultry litter biochar. *Appl. Energy* **236**, 1269–1279 (2019).
16. Dessi, P. et al. Carboxylic acids production and electrosynthetic microbial community evolution under different $CO_2$ feeding regimens. *Bioelectrochemistry* **137**, 107686 (2021).
17. Liu, R., Hao, X. & Wei, J. Function of homoacetogenesis on the heterotrophic methane production with exogenous $H_2/CO_2$ involved. *Chem. Eng. J.* **284**, 1196–1203 (2016).
18. Raes, S. M. T., Jourdin, L., Buisman, C. J. N. & Strik, D. P. B. T. B. Continuous long-term bioelectrochemical chain elongation to butyrate. *Chem. Electro Chem.* **4**, 386–395 (2017).
19. Gophna, U., Konikoff, T. & Nielsen, H. B. *Oscillospira* and related bacteria - From metagenomic species to metabolic features. *Environ. Microbiol.* **19**, 835–841 (2017).
20. Isipato, M. et al. Propionate production by bioelectrochemically-assisted lactate fermentation and simultaneous $CO_2$ recycling. *Front. Microbiol.* **11**, 599438 (2020).
21. Dong, J. & Horvath, S. Understanding network concepts in modules. *BMC Syst. Biol.* **1**, 1–20 (2007).
22. Ma, H. W. & Zeng, A. P. The connectivity structure, giant strong component and centrality of metabolic networks. *Bioinformatics* **19**, 1423–1430 (2003).
23. Ma, B. et al. Earth microbial co-occurrence network reveals interconnection pattern across microbiomes. *Microbiome* **8**, 1–12 (2020).
24. Ao, T. et al. Anaerobic thermophilic digestion of maotai-flavored distiller's grains: Process performance and microbial community dynamics. *Energy Fuels* **33**, 8804–8811 (2019).
25. Stevenson, B. S. et al. Microbial communities in bulk fluids and biofilms of an oil facility have similar composition but different structure. *Environ. Microbiol* **13**, 1078–1090 (2011).
26. Zhang, D. et al. Substantially enhanced anaerobic reduction of nitrobenzene by biochar stabilized sulfide-modified nanoscale zero-valent iron: Process and mechanisms. *Environ. Int.* **131**, 105020 (2019).
27. Zhang, X. et al. *Petroclostridium xylanilyticum* gen. nov., sp. nov., a xylan-degrading bacterium isolated from an oilfield, and reclassification of clostridial cluster III members into four novel genera in a new *Hungateiclostridiaceae fam*. nov. *Int. J. Syst. Evol. Microbiol* **68**, 3197–3211 (2018).
28. Pedros-Alio, C. Marine microbial diversity: can it be determined? *Trends Microbiol.* **14**, 257–263 (2006).
29. Saunders, A. M., Albertsen, M., Vollertsen, J. & Nielsen, P. H. The activated sludge ecosystem contains a core community of abundant organisms. *ISME J.* **10**, 11–20 (2016).
30. Zhao, R., Liu, J., Feng, J., Li, X. & Li, B. Microbial community composition and metabolic functions in landfill leachate from different landfills of China. *Sci. Total Environ.* **767**, 144861 (2021).
31. Ping, Q., Zheng, M., Dai, X. & Li, Y. Metagenomic characterization of the enhanced performance of anaerobic fermentation of waste activated sludge with $CaO_2$ addition at ambient temperature: Fatty acid biosynthesis metabolic pathway and CAZymes. *Water Res.* **170**, 115309 (2020).
32. Atsumi, S., Hanai, T. & Liao, J. C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86–89 (2008).
33. Siebers, B. et al. Reconstruction of the central carbohydrate metabolism of *Thermoproteus tenax* by use of genomic and biochemical data. *J. Bacteriol.* **186**, 2179–2194 (2004).
34. Kanao, T., Fukui, T., Atomi, H. & Imanaka, T. ATP-citrate lyase from the green sulfur bacterium *Chlorobium limicola* is a heteromeric enzyme composed of two distinct gene products. *Eur. J. Biochem.* **268**, 1670–1678 (2001).
35. Charubin, K. & Papoutsakis, E. T. Direct cell-to-cell exchange of matter in a synthetic Clostridium syntrophy enables $CO_2$ fixation, superior metabolite yields, and an expanded metabolic space. *Metab. Eng.* **52**, 9–19 (2019).
36. Fast, A. G. & Papoutsakis, E. T. Stoichiometric and energetic analyses of non-photosynthetic $CO_2$-fixation pathways to support synthetic biology strategies for production of fuels and chemicals. *Curr. Opin. Chem. Eng.* **1**, 380–395 (2012).
37. Battin, T. J., Besemer, K., Bengtsson, M. M., Romani, A. M. & Packmann, A. I. The ecology and biogeochemistry of stream biofilms. *Nat. Rev. Microbiol.* **14**, 251–263 (2016).
38. Werner, J. J. et al. Bacterial community structures are unique and resilient in full-scale bioenergy systems. *Proc. Natl Acad. Sci. U.S.A.* **108**, 4158–4163 (2011).
39. Patil, S. A. et al. Selective enrichment establishes a stable performing community for microbial electrosynthesis of acetate from $CO_2$. *Environ. Sci. Technol.* **49**, 8833–8843 (2015).
40. Lin, R. et al. Boosting biomethane yield and production rate with graphene: The potential of direct interspecies electron transfer in anaerobic digestion. *Bioresour. Technol.* **239**, 345–352 (2017).
41. Marshall, C. W. et al. Metabolic reconstruction and modeling microbial electrosynthesis. *Sci. Rep.* **7**, 8391 (2017).
42. Denger, K. & Schink, B. Energy conservation by succinate decarboxylation in Veillonella parvula. *Microbiology* **138**, 967–971 (1992).
43. Castilla-Archilla, J. et al. Screening for suitable mixed microbial consortia from anaerobic sludge and animal dungs for biodegradation of brewery spent grain. *Biomass. Bioenergy* **159**, 106396 (2022).
44. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
45. Callahan, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
46. Douglas, G. M. et al. PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* **38**, 685–688 (2020).
47. Caspi, R. et al. The MetaCyc database of metabolic pathways and enzymes—a 2019 update. *Nucleic Acids Res.* **48**, 445–453 (2020).
48. McKight, P. E. & Najab, J. Kruskal-Wallis test. *The Corsini Encyclopedia of Psychology*, 1–1 (John Wiley & Sons, 2010).
49. Parks, D. H., Tyson, G. W., Hugenholtz, P. & Beiko, R. G. STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* **30**, 3123–3124 (2014).
50. Lepš, J. & Šmilauer, P. Multivariate analysis of ecological data using CANOCO (Cambridge University Press, 2003).
51. Li, M., Li, D., Tang, Y., Wu, F. & Wang, J. CytoCluster: a cytoscape plugin for cluster analysis and visualization of biological networks. *Int. J. Mol. Sci.* **18**, 1880 (2017).

## ACKNOWLEDGEMENTS

The authors thank Yaxue He, Juan Castilla-Archilla, Leah Egan, Marco Isipato, and Paolo Dessì, (NUIG, Ireland) for their help and support during the data-collection work. This publication has emanated from research supported by Science Foundation Ireland (SFI) through the SFI Research Professorship Program entitled *Innovative*

*Energy Technologies for Biofuels, Bioenergy and a Sustainable Irish Bioeconomy* (IETSBIO[3]; grant number 15/RP/2763) and the Research Infrastructure research grant *Platform for Biofuel Analysis* (Grant Number 16/RI/3401).

## AUTHOR CONTRIBUTIONS

Q.Y. writing-original draft preparation and editing, data analysis, and investigation. G.W. draft editing, validation, and methodology. P.N.L.L. draft editing, conceptualization, and supervision.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41545-022-00180-3.

**Correspondence** and requests for materials should be addressed to Qidong Yin.