Environmental Science & Technology

pubs.acs.org/est

Article

Functional Redundancy Secures Resilience of Chain Elongation Communities upon pH Shifts in Closed Bioreactor Ecosystems

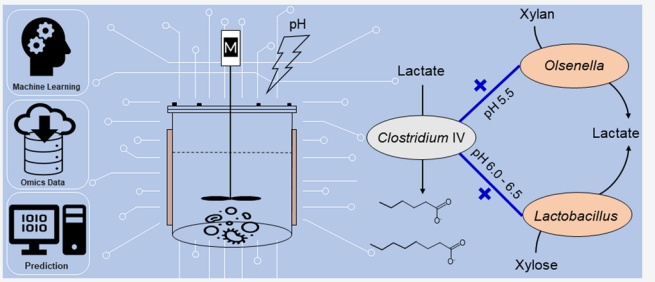
Bin Liu, Heike Sträuber, Florian Centler, Hauke Harms, Ulisses Nunes da Rocha, and Sabine Kleinsteuber*

Cite This: *Environ. Sci. Technol.* 2023, 57, 18350−18361

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: For anaerobic mixed cultures performing microbial chain elongation, it is unclear how pH alterations affect the abundance of key players, microbial interactions, and community functioning in terms of medium-chain carboxylate yields. We explored pH effects on mixed cultures enriched in continuous anaerobic bioreactors representing closed model ecosystems. Gradual pH increase from 5.5 to 6.5 induced dramatic shifts in community composition, whereas product range and yields returned to previous states after transient fluctuations. To understand community responses to pH perturbations over long-term reactor operation, we applied Aitchison PCA clustering, linear mixed-effects models, and random forest classification on 16S rRNA gene amplicon sequencing and process data. Different pH preferences of two key chain elongation species—one *Clostridium* IV species related to *Ruminococcaceae* bacterium CPB6 and one *Clostridium sensu stricto* species related to *Clostridium luticellarii*—were determined. Network analysis revealed positive correlations of *Clostridium* IV with lactic acid bacteria, which switched from *Olsenella* to *Lactobacillus* along the pH increase, illustrating the plasticity of the food web in chain elongation communities. Despite long-term cultivation in closed systems over the pH shift experiment, the communities retained functional redundancy in fermentation pathways, reflected by the emergence of rare species and concomitant recovery of chain elongation functions.



KEYWORDS: *carboxylate platform, medium-chain carboxylates, lactate-based chain elongation, reactor microbiome, time series analysis, compositional data, machine learning*

INTRODUCTION

Microbial ecologists aim to understand the main environmental factors driving the processes of microbial community assembly and functioning.[1−3] Ecological selection exerted by abiotic and biotic factors influences the growth rates of community members and their interactions, thereby determining the composition and functioning of microbial communities.[4−7] In engineered systems, pH is a key parameter shaping microbial communities and steering them toward specific functions.[8−12]

To produce platform chemicals such as *n*-butyrate (C4), *n*-caproate (C6), and *n*-caprylate (C8) from renewable resources sustainably, lactate-based microbial chain elongation (CE) coupled with *in situ* lactate formation holds promise to valorize organic waste streams or biomass residues within the carboxylate platform.[13] Efficient and stable CE processes rely on trophic interactions in microbial communities with complementary and parallel metabolic functions in a food web.[14] In this context, pH can substantially affect different cooperating or competing community members and hence the resulting product profile of CE processes. For example, Candry et al.[15] found that pH values below 6 favored the production of C6 over that of propionate, whereas a CE community adapted to pH 5.5 shifted to propionate as the dominant product upon a pH shift to 6.5. In another open CE system fed with lactate-rich silage, pH variation induced the development of distinct key subcommunities, reflecting different pH optima for the production of C6 and C8.[16] Due to the complex dynamics of open mixed cultures with continuous reinoculation and undefined feedstocks, it is challenging to understand the impact of pH on key players of CE, microbial interactions, and product profiles in a quantifiable and predictable way. To overcome such obstacles, we previously established model CE ecosystems using anaerobic bioreactors without continued inoculation and with xylan and lactate as model substrates to simulate the feedstock conditions in anaerobic fermentation fed with ensiled plant biomass.[14] The reactors produced C4, C6, and C8 from xylan and lactate, and

Special Issue: Data Science for Advancing Environmental Science, Engineering, and Technology

Received: December 20, 2022
Revised: April 13, 2023
Accepted: April 13, 2023
Published: April 25, 2023

cooperation as well as competition between different functional groups established under constant process conditions.[14] We further demonstrated that shortening the hydraulic retention time shapes CE communities toward desired C6 and C8 production.[17] Here, we explored the effect of pH shifts on CE communities by benefiting from these previously established model ecosystems.

Even without continued inoculation, such closed systems are relatively complex regarding microbial interactions and metabolic processes. Enrichment cultures can maintain their functional stability by self-assembly, which appears challenging for designing synthetic communities, due to the lack of knowledge required to rationally engineer stable microbial interactions.[18] Next-generation sequencing (e.g., 16S rRNA amplicon sequencing) allows for capturing the dynamics of entire communities with high phylogenetic resolution over long-term experiments,[7] although there are some methodological limitations, such as PCR biases.[19] Additionally, amplicon sequencing data (e.g., amplicon sequencing variants − ASV) only provide proportions. Considering the compositionality of such data sets that contain the relationship information between the parts, approaches that usually start with a log-ratio transformation were developed to avoid the common pitfalls in analyzing compositional data.[20−22] For correlation analysis, association network algorithms are commonly applied, inferring nonrandom co-occurrence patterns between community members and assessing microbial responses to environmental changes. In this study, standard microbiome analysis and compositional data analysis were implemented to achieve statistically robust results.

Besides pH, time is an essential component in long-term experimental studies. We categorized time as another factor to emphasize the effect of time (of pH shift) on community dynamics, which reflects the ecological memory in our ecosystems. Suitably clocked sampling with replicates over long experimental times gives insight into the stability of microbial communities and their response to and recovery from perturbations.[9,23,24] Linear mixed-effects models (LME) and variations thereof are commonly used for modeling time-resolved 16S rRNA amplicon sequencing data, thereby identifying temporal microbial interaction patterns.[25,26] We hypothesized that the pH value predominantly determines the assembly of CE reactor microbiomes, but the impact of time needs to be disentangled by applying LME. The temporal patterns of identified taxa are crucial to understand their roles in CE functions being inferred from the correlation with measured process parameters. Feature selection using random forest classification was performed to denote bioindicators of pH changes. Subsequently, the genetic potential of these bioindicators was investigated by functional annotation of metagenome-assembled genomes (MAGs).[17] As for CE, it is still unclear how the different microorganisms interact and what conditions they thrive in. In this context, pH can be a critical parameter that affects these relationships and ultimately the end products of CE. Our study focused on the effects of pH increase considering three aspects: (i) the identity and abundance of key players of lactate-based CE, (ii) the effect on microbial interactions, and (iii) the functional resilience of the CE reactor microbiome. Understanding the underlying ecological principles of CE reactor microbiomes is the foundation for the development of more efficient and stable mixed-culture bioprocesses within the framework of green chemistry and a sustainable circular economy.

MATERIALS AND METHODS

Reactor Operation and Sampling. A microbial community was enriched in a 1 L bioreactor (BIOSTAT A plus, Sartorius, Göttingen, Germany), inoculated with broth from a former study,[16] and fed with mineral medium containing xylan and lactate over 150 days.[14] The enriched community producing C4, C6, and C8 was further selected by reducing the hydraulic retention time in two parallel BIOSTAT bioreactors (A and B) for almost one year.[17] The pH was kept at 5.5 in both periods. Here, we tested the effect of pH increase with a fixed retention time of 4 days. Before starting the experiment, the microbial communities of bioreactors A and B were equally distributed by pumping the content from A to B and back while maintaining anoxic conditions.

The reactor configuration was similar as reported before,[14] with both bioreactors operated at 38 ± 1 °C, constantly stirred at 150 rpm, and the pH automatically controlled with 5 M NaOH. For daily feeding, 2.94 g of lactate and 2.50 g of water-soluble xylan were supplied in 0.25 L of anoxic mineral medium (composition as described previously[14]). The same volume of completely mixed effluent was harvested daily from the reactors before feeding. The starting pH was 5.5 for both bioreactors. After 42 days, we increased the pH of bioreactor A to 6.0 and further to 6.5 from day 112 to day 238. To consider the effect of time on community assembly, a different temporal scheme of pH increase was applied in reactor B (pH 5.5, days 0−144; pH 6.0, days 145−214; pH 6.5, days 215−238).

Reactor headspace and effluent were sampled twice per week. In total, 68 samples were collected from each reactor during 238 days of operation. The effluent was centrifuged, and the supernatant was used for measuring concentrations of xylan, carboxylates, and alcohols.[14] Optical density (OD) at 600 nm of the effluent was measured before centrifugation. Pelleted cells were stored at −20 °C for DNA-based community analysis.[14]

Analytical Methods. Daily gas production was monitored as described previously.[27] Gas composition was determined in triplicate for H_2, CO_2, N_2, CH_4, and O_2 by gas chromatography.[28] Concentrations of carboxylates and alcohols were analyzed in triplicate by gas chromatography, and xylan was measured by a modified dinitrosalicylic acid reagent method.[14] At the beginning and the end of the experiment, cell mass concentration was calculated from OD values correlated with cell dry mass,[14] with mean correlation coefficients of 1 OD_{600} = 0.641 g L^{-1} for bioreactor A and 1 OD_{600} = 0.632 g L^{-1} for bioreactor B.

Total DNA was isolated from frozen cell pellets using a NucleoSpin Microbial DNA Kit (Macherey-Nagel, Düren, Germany). Methods for DNA quality control and quantification were reported previously.[29] 16S rRNA genes were PCR-amplified using primers 341f and 785r[30] and sequenced on the Illumina MiSeq platform (MiSeq Reagent Kit v3, 2 × 300 bp; Illumina, San Diego, CA) according to the MiSeq manual.

Microbiome Data Analysis. The QIIME 2 v2020.2 pipeline[31] with DADA2 plugin[32] was applied for demultiplexing sequences, filtering phiX reads, denoising, merging read pairs, trimming, and removing chimeras. A total of 6,855,572 sequences ranging from 21,437 to 66,272 read pairs per sample were obtained, with a median of 50,439 in 136 samples. A feature table was created indicating the frequency of each ASV clustered at 100% identity. ASVs with frequencies >2 in at least three samples were kept for further analyses. Taxonomy was assigned with a naïve Bayes classifier trained on the database MiDAS

https://doi.org/10.1021/acs.est.2c09573
Environ. Sci. Technol. 2023, 57, 18350−18361

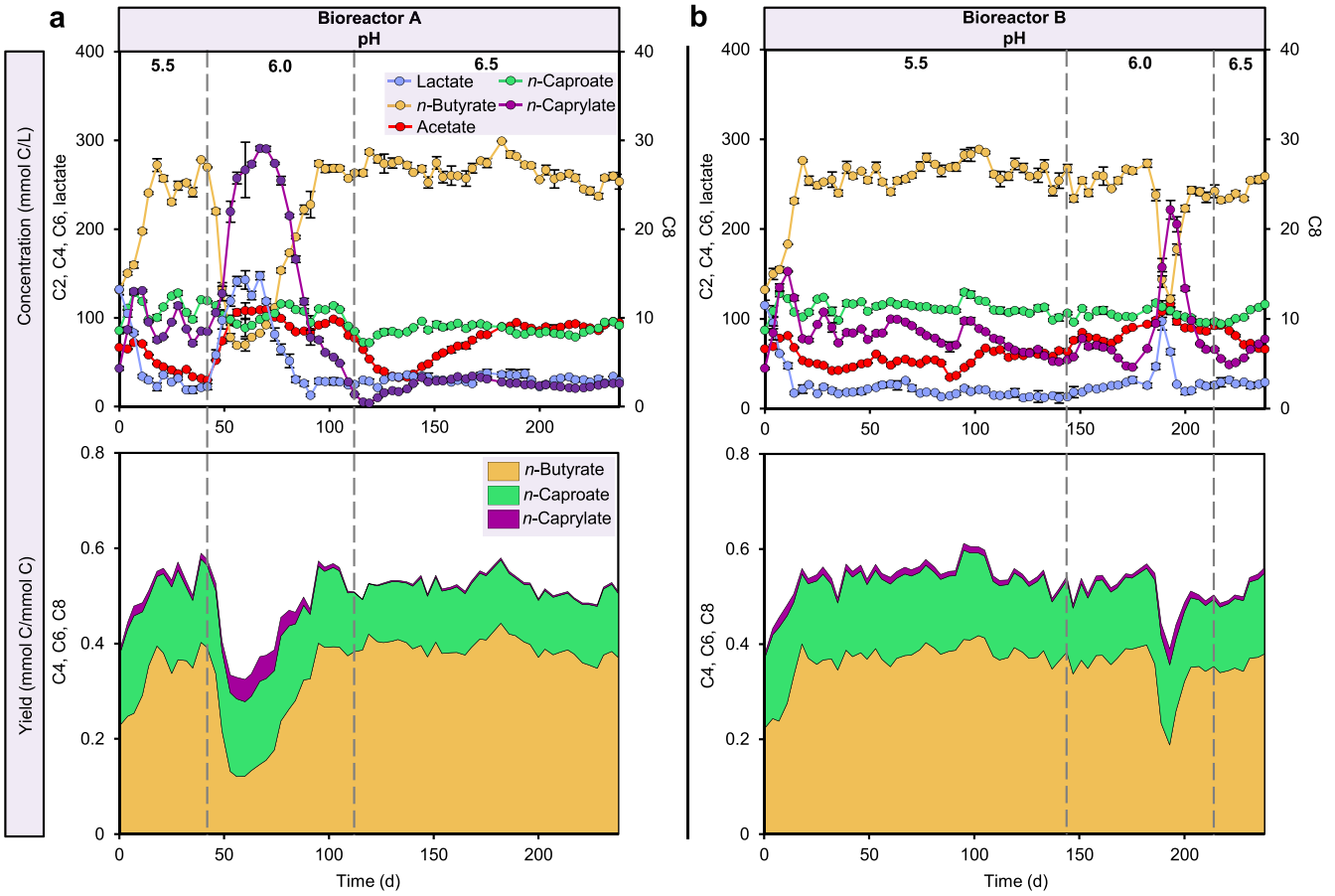


Figure 1. Performance of bioreactors. Concentrations of fermentation products and lactate, as well as yields of chain elongation products in bioreactors (a) A and (b) B at three pH levels. Yield is given in C mole of product to substrate ratio. Fermentation products: C2, acetate; C4, *n*-butyrate; C6, *n*-caproate; C8, *n*-caprylate.

2.1[33] and curated with the RDP Classifier 2.2[34] (confidence threshold: 80%). The filtered ASV table (see the Supporting Information) was rarefied to 21,389 reads for downstream analyses (rarefaction curves reached the plateau, Figure S1). Following the common practice to normalize samples to the smallest sample size,[35,36] we did not intend to subjectively discard any samples, which may cause difficulties for downstream analyses. As the microbial communities were highly enriched, we assumed that 21,389 reads are sufficient to cover most ASVs. A total of 97 unique ASVs remained after rarefaction.

α-Diversity based on rarefied ASV data was evaluated by calculating diversity, evenness, and richness.[37] The indices of order one (^{1}D and ^{1}E) quantify the diversity and evenness by weighting all ASVs equally, whereas the indices of order two (^{2}D and ^{2}E) give more weight to the dominant ASVs. Considering the compositional nature of amplicon sequencing data,[19] we analyzed the data with standard approaches and their compositional replacements. For dissimilarities in community composition (β-diversity), we used Bray−Curtis distance-based principle coordinate analysis (PCoA)[38] and Aitchison principal component analysis (PCA) via DEICODE, which is robust to data sparsity.[20] Compared to Bray−Curtis, Aitchison PCA using DEICODE solves the problems of high sparsity of 16S rRNA amplicon sequencing data via two steps: a compositional processing using the centered log-ratio transform on nonzero values of the data and a reduction of dimensionality through robust PCA on those nonzero values. The QIIME 2 plugin Qurro[39] was used to visualize and explore feature rankings in the produced DEICODE biplot. PERMANOVA ("adonis" function in R vegan package, v2.5.6; 999 permutations)[21] was used for statistical analyses of β-diversity, with P values adjusted according to the false discovery rate controlling procedure introduced by Benjamini and Hochberg.[40] The metagenome data included in this study and a detailed description of MAG reconstruction can be found in our previous study.[17]

Statistical Analysis of Effects of pH Increase on Reactor Microbiota Time Series. A redundancy analysis-based variation partitioning analysis (VPA) was used to quantify the relative contribution of individual process parameters (pH and time) and their interactive effects on temporal variation in community composition. VPA was performed using the "varpart" function in the R package vegan. We performed a partial Mantel test for each process parameter to examine its correlation with community composition represented by Aitchison and Bray−Curtis distances, independent of time (9999 permutations) using vegan.

The QIIME 2 plugin q2-longitudinal with default settings was used to construct the LME for regression analyses involving dependent data.[26] Random intercept models (REML method) were used to track longitudinal changes of metrics including α- and β-diversity and ASV abundances. In brief, pH and time were designated as fixed effects and bioreactor as a random effect, whereas values represent samples of a random collection. The

https://doi.org/10.1021/acs.est.2c09573
Environ. Sci. Technol. 2023, 57, 18350−18361

response variables are the following metrics: ^{1}D, ^{2}D, ^{1}E, ^{2}E, richness, PC1 of Aitchison or Bray−Curtis, and ASV abundance.

The microbial temporal variability linear mixed model (MTV-LMM) was used to identify autoregressive taxa and predict their relative abundances at later time points.[25] The model assumes that the temporal changes in relative abundance of ASVs are a time-homogeneous high-order Markov process. To select the core time-dependent taxa, MTV-LMM was applied to each individual pH level, which generated a temporal kinship matrix representing the similarity between every pair of normalized ASV abundances (a given time for a given individual) across time. A concept of time-explainability was introduced to quantify the temporal variance explained by the microbial community at previous time points.

Random Forest (RF) Classification. Supervised classification of pH levels on community compositions was performed using QIIME 2 q2-sample-classifier with default settings.[41] Rarefied ASV data were used as features to train and test the classifier. First, a nested cross-validation of the RF model was applied to overview the classification of the pH levels for all samples. For model optimization, a second layer of cross validation (outer loop) was incorporated to split the data set into training and test sets five times, and therefore, each sample ended up in a test set once. During each iteration of the outer loop, the training set is split again five times in an inner loop to optimize parameter settings for estimation of that fold. Five different final models were trained, with each sample receiving a predicted value. The overall accuracy was calculated by comparing the predicted values to the true values.

Next, we performed a feature selection by randomly picking 80% of the samples to train an RF classifier, and the remaining 20% of the samples were used to test the classification accuracy of the classifier. K-fold cross-validation (K = 5) was performed during automatic feature selection and parameter optimization steps to tune the model. As determined by using recursive feature elimination, the most important features that maximized model accuracy were selected. Model accuracy and predictions were based on the classifier that utilized the reduced feature set.

Network Analysis. Co-occurrence networks based on rarefied ASV data and process parameter data were inferred by using FlashWeave v0.16 implemented in Julia.[22] FlashWeave uses the centered log-ratio approach for the correction of compositional microbial abundances and infers direct associations. Three networks were constructed for the three individual pH levels, which featured a correlation coefficient <−0.5 or >0.5. Another network was constructed from the entire data of all pH levels. All networks were visualized in Cytoscape v3.8.0.[42]

RESULTS

Fluctuation and Recovery of Process Performance. The pH increase from 5.5 to 6.0 caused fluctuations in fermentation products and lactate concentrations, which were not observed upon further increase to 6.5 (Figure 1). First, we applied the pH increase in bioreactor A, which immediately presented an increased C8 concentration up to 29.1 mmol C/L, corresponding to a yield (C mole product to substrate ratio) of 5.2, and a relatively stable yield of C6 (16.0 ± 1.5 at pH 6.0). Lactate and acetate accumulated to concentrations of 147.5 and 109.7 mmol C/L, respectively; while C4 concentration dropped to 69.1 mmol C/L, with a yield of 12.1 (Figure 1a). The pH increase left the fast consumption of xylan unaffected (Figure S2). Afterward, accumulated lactate and acetate were consumed and C4 concentration returned to the previous level with 273.9 mmol C/L on day 95 at pH 6.0. Notably, further pH increase to 6.5 did not result in such fluctuations (Figure 1a). Later, we replicated the pH increase from 5.5 to 6.5 in bioreactor B to confirm the observed effects of pH increase. With longer operation at pH 5.5 for 144 days, comparable fluctuations in concentrations of lactate, acetate, C4 and C8 were observed, but with a delay of 38 days after the pH increase to 6.0. Concentrations of lactate, acetate, C4, C6, and C8 were relatively stable when bioreactor B was operated at pH 6.5. The pH increase also resulted in fluctuations of daily gas production and gas composition (Figure S3). A general upward trend of cell mass yield at pH 6.5 suggests a facilitating effect of higher pH on the growth of enriched populations (Figure S4).

Microbial Community Shifts and Emergence of Rare Species. After the pH increase, α-diversity metrics showed decreases in diversity (^{1}D) and evenness (^{1}E) but an increase in richness (Figure 2; similar results for ^{2}D and ^{2}E shown in Figure S5). We used LME models to test whether these indices were impacted by pH and time, which presents the memory effect on community dynamics. Three separate LME models were fitted to examine ^{1}D, ^{1}E, and richness across pH gradients because the trajectories appeared nearly linear. Diversity was significantly impacted by pH (P < 0.001) and time (P < 0.001), indicating

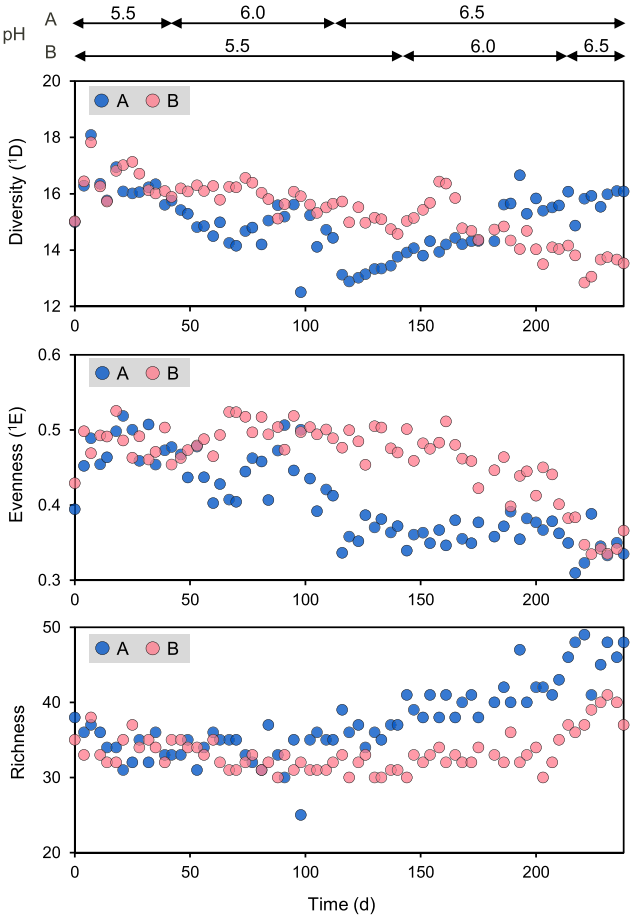


Figure 2. Longitudinal changes in α-diversity at three pH levels. Based on the relative abundances of ASVs, we calculated the α-diversity represented by (a) diversity of order one (^{1}D), (b) evenness of order one (^{1}E), and (c) richness. Diversity and evenness of order one were quantified by weighting all ASVs equally. A and B stand for bioreactors A and B.

https://doi.org/10.1021/acs.est.2c09573
Environ. Sci. Technol. 2023, 57, 18350−18361

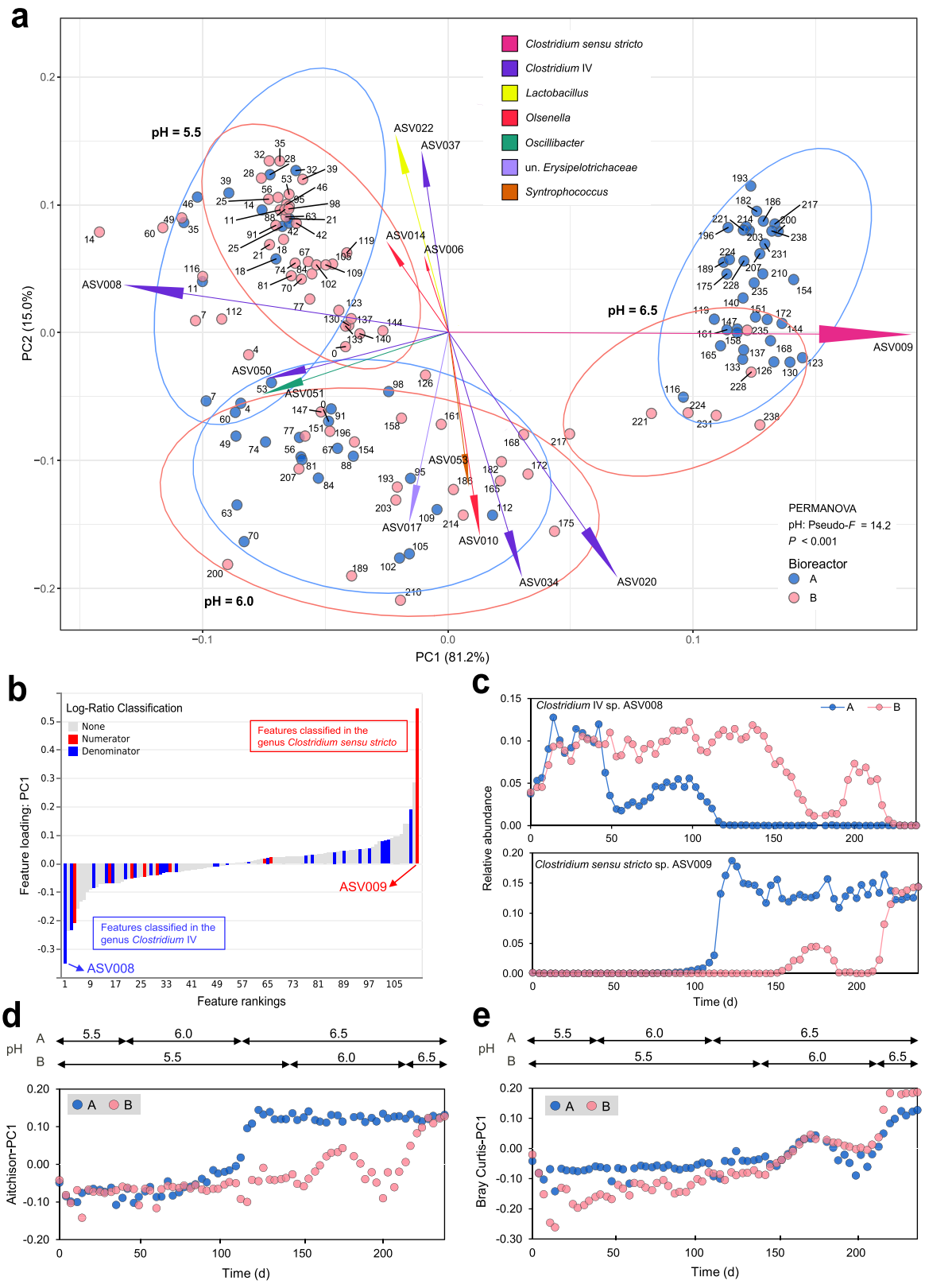


Figure 3. Effects of pH increase and time on bacterial community composition. (a) Variance-based compositional principal component analysis (PCA) biplot based on Aitchison distance. Dots are named according to sampling days. Ellipses of 95% confidence intervals were added to each individual pH level of the bioreactors. The size of an ASV arrow indicates the strength of the relationship of that ASV to the community composition. ASVs are colored by family. (b) ASV ranks estimated from Aitchison distance-based PCA (PC1) with *Clostridium* IV and *Clostridium sensu stricto* highlighted. (c) Longitudinal changes in relative abundances of *Clostridium* IV sp. ASV008 and *Clostridium sensu stricto* sp. ASV009 at three pH levels. (d and e) Longitudinal changes in β-diversity at three pH levels, based on Aitchison (d) and Bray−Curtis (e) dissimilarities. A and B stand for bioreactors A and B. un., unclassified.

https://doi.org/10.1021/acs.est.2c09573
Environ. Sci. Technol. 2023, 57, 18350−18361

that diversity was reduced much stronger by pH with a factor of 6.188 than by time with a factor of 0.209 (Table S1). Evenness and richness were also significantly associated with pH and time, although pH exerted much stronger impacts on both indices (Tables S2 and S3). As shown in Figure S6, the relative ASV abundances categorized from phylum to genus level varied along the pH gradients, e.g., *Actinomyces* and *Prevotella* became apparent at pH 6.5 along with an increasing abundance of *Clostridium sensu stricto* and decreasing abundances of *Clostridium* IV and *Eubacterium* (Figure S6e).

β-Diversity analysis revealed that the bacterial communities differed significantly between the three pH levels (PERMANOVA; $P < 0.001$) (Figure 3a and Figure S7). ASVs of *Clostridium* IV, *Oscillibacter*, *Olsenella*, and *Syntrophococcus* were strongly associated with the communities at pH 5.5 and 6.0, whereas *Clostridium sensu stricto* ASV009 was most strongly associated with the communities at pH 6.5 (Figure 3a). Based on the association with dissimilarities in community composition, *Clostridium* IV ASV008 (lowest ranked taxon) and *Clostridium sensu stricto* ASV009 (highest ranked taxon) correspond to the most influential taxa driving the Aitchison PCA clustering (Figure 3b). After fitting LME models to their dynamics in relative abundance (Figure 3c), results showed that the relative abundance of ASV008 was significantly impacted by pH ($P < 0.001$) and time ($P = 0.002$), whereas only pH ($P < 0.001$) significantly impacted the abundance of ASV009 (Tables S4 and S5). In both cases, pH had a much stronger impact than time. By applying LME models, we examined how *β*-diversity changed over time in each bioreactor (Figure 3d,e). Results indicated that pH was the most influencing factor, although time had significant effects as well (Tables S6 and S7). To understand the impact of only pH on the community assembly, we removed the effect of time using partial Mantel tests. We correlated the time-corrected dissimilarities of community composition with pH, and the results show strong, significant correlations based on Aitchison distance ($r_m = 0.61$, $P < 0.001$) and Bray−Curtis distance ($r_m = 0.72$, $P < 0.001$) (Table 1). We further considered the impact of pH and time in a quantifiable way using VPA. Evaluation of the overall contributions of pH and time indicated that together they explain 61% of the microbial community variations based on Bray−Curtis (Figure S8), which also reflects that additional factors such as stochastic assembly processes or chemical effects of CE products played a role. In total, 24% and 3% of the variations were independently explained by pH and time, respectively. These results support those inferred from the LME models.

pH Bioindicators and Time-Dependent Taxa. Overall, the nested cross-validation of the RF model represented a classification accuracy of 97.8% in matching the predicted three pH levels (5.5, 6.0, and 6.5) with the true pH levels for all 136 samples (Figure S9), using ASV data to follow community composition dynamics. We performed recursive feature elimination with cross-validation; the 18 most important features were selected that gave perfect discrimination between the three pH levels (Figure 4). These ASVs were defined as pH bioindicators, belonging to the genera *Clostridium* IV, *Syntrophococcus*, *Lactobacillus*, *Olsenella*, *Bulleidia*, *Clostridium sensu stricto*, *Eubacterium*, *Lachnospiraceae incertae sedis*, *Sporanaerobacter*, and *Actinomyces* (Figure 4b). Among these pH bioindicators, four increased in abundance while 14 became less abundant along the pH increase. Notably, the most influential ASVs driving the Aitchison PCA clustering were also pH bioindicators, including the abundant taxa *Clostridium* IV ASV008 and *Clostridium sensu stricto* ASV009 (Figure 4b).

Table 1. Partial Mantel Tests Showing Significant Correlations between the Time-Corrected Dissimilarities of Microbial Community Composition and Process Parameters

	Aitchison distance		Bray−Curtis distance	
process parameter	r_m [a]	P [b]	r_m	P
pH	0.61	<0.001	0.72	<0.001
Conc. C2 [c]	0.27	<0.001	0.18	<0.001
Conc. C4	0.07	0.013	−0.01	0.569
Conc. C6	0.29	<0.001	0.48	<0.001
Conc. C8	0.25	<0.001	0.16	<0.001
Conc. lactate	0.02	0.258	0.01	0.401
Conc. biomass	0.16	<0.001	0.11	0.002
yield C2	0.27	<0.001	0.15	<0.001
yield C4	0.09	0.004	0.00	0.448
yield C6	0.38	<0.001	0.40	<0.001
yield C8	0.22	<0.001	0.13	0.003
yield biomass	0.09	0.001	0.06	0.037
O_2	0.43	<0.001	0.44	<0.001
CO_2	0.14	<0.001	0.18	<0.001
H_2	0.19	<0.001	0.16	<0.001
time	0.14	<0.001	0.33	<0.001

[a] r_m, the correlation coefficient based on partial Mantel test, in which time was controlled. The permutation test compares the original r_m to r_m computed in 9999 random permutations. [b] The reported P value is one-tailed. [c] Conc., concentration

As pH and time were the two most influencing factors of microbial community assembly, we intended to disentangle the ASVs that were mostly associated with time rather than with pH. By using MTV-LMM, we identified time-dependent (autoregressive) taxa, whose abundance can be predicted based on the previous community composition. They reflect the ecological memory effect on community dynamics, i.e., that past events influence the present trajectory of community composition. In this longitudinal study, 32, 25, and 40 ASVs were predicted to be significantly ($P < 0.05$) affected by the past composition of the community at pH 5.5, 6.0 and 6.5, respectively, with the time-explainability ranging from 17% to 80%, 17% to 83% and 13% to 96%, respectively (Figure S10).

Microbial Interaction Patterns. Partial Mantel test showed significant correlations of the community composition with process performance and the changing conditions (Table 1). Consequently, we constructed an overall network and three separate networks for each pH level to discern the succession of microbial interactions and reveal potential metabolic functions. After the pH increase to 6.5, more nodes and edges and higher average clustering coefficient and heterogeneity were found, suggesting that the overall interaction intensity was higher at pH 6.5 (Table S8). In agreement with Aitchison PCA analysis, pH was significantly correlated with pH bioindicators ASV008 and ASV009 (Figure S11). Changes of interaction patterns over pH are shown in Figure 5. At the family level, *Ruminococcaceae* co-occurred with *Lachnospiraceae* and *Erysipelotrichaceae* at all pH levels, while it co-occurred with *Coriobacteriaceae* only at pH 5.5. *Ruminococcaceae* also co-occurred with *Lactobacillaceae* at pH 6.0 and 6.5 and with *Actinomycetaceae* only at pH 6.5. *Clostridiaceae* 1 co-occurred with *Clostridiales incertae sedis* XI and *Erysipelotrichaceae* only after the pH increase to 6.0. *Erysipelotrichaceae* showed positive correlations with *Lactobacillaceae* at pH 6.0 and 6.5, where its negative correlation with

https://doi.org/10.1021/acs.est.2c09573
Environ. Sci. Technol. 2023, 57, 18350−18361

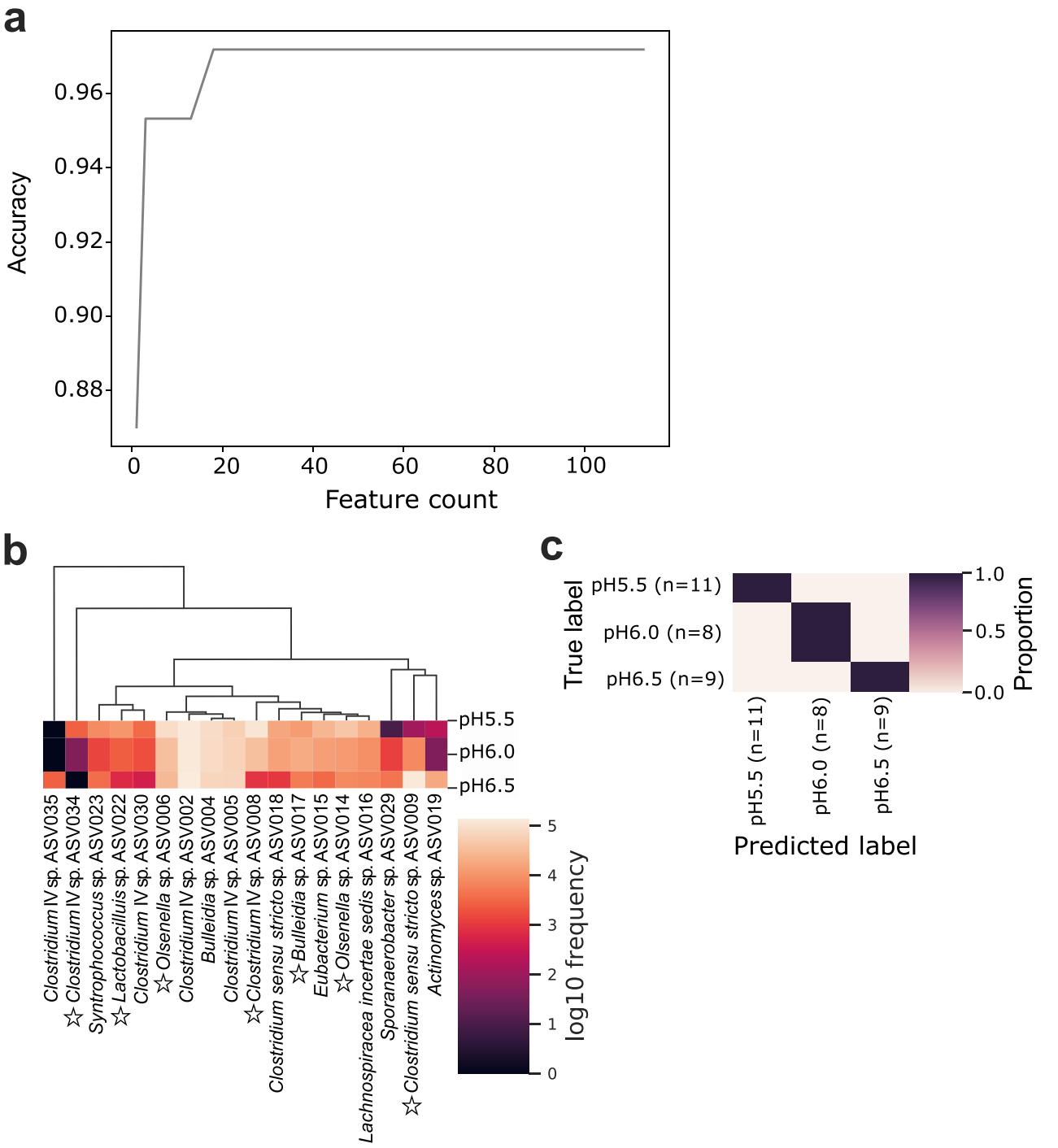


Figure 4. pH bioindicators determined by random forest classification accurately predict the different pH levels. (a) Recursive feature elimination plot illustrating the model accuracy changes as a function of ASV count. The top-ranked 18 ASVs (pH bioindicators) that maximize accuracy are automatically selected for optimizing the model, based on their mean decrease in Gini scores, according to their ASV abundance distribution, with pH as the response variable. (b) Heatmap showing dynamics of the mean abundance of pH bioindicators at the different pH levels. ASVs shown in Aitchison PCA biplot are indicated by a star. (c) Confusion matrix for the optimal classifier of samples at different pH levels. The classifier was trained on the randomly picked 80% of the samples, which was then tested on the remaining 20%. Overall accuracy was calculated by comparing the predicted values with the true values.

Coriobacteriaceae vanished. Notably, the positive correlation of *Erysipelotrichaceae* with *Lachnospiraceae* was not seen at pH 6.0. The positive correlation between C6 yield and *Eubacterium* ASV015 was presented in the overall network and the individual networks of pH 5.5 and pH 6.0 but not in that of pH 6.5 (Figure S11 and Figure 5). In general, stronger correlations ($|r| > 0.5$) were observed at pH 6.5, including the negative correlation of *Prevotella* ASV041 with *Bulleidia* ASV017.

DISCUSSION

Different pH Niches of Chain Elongation Key Players *Clostridium* IV and *Clostridium sensu stricto*. The

https://doi.org/10.1021/acs.est.2c09573
Environ. Sci. Technol. 2023, 57, 18350−18361

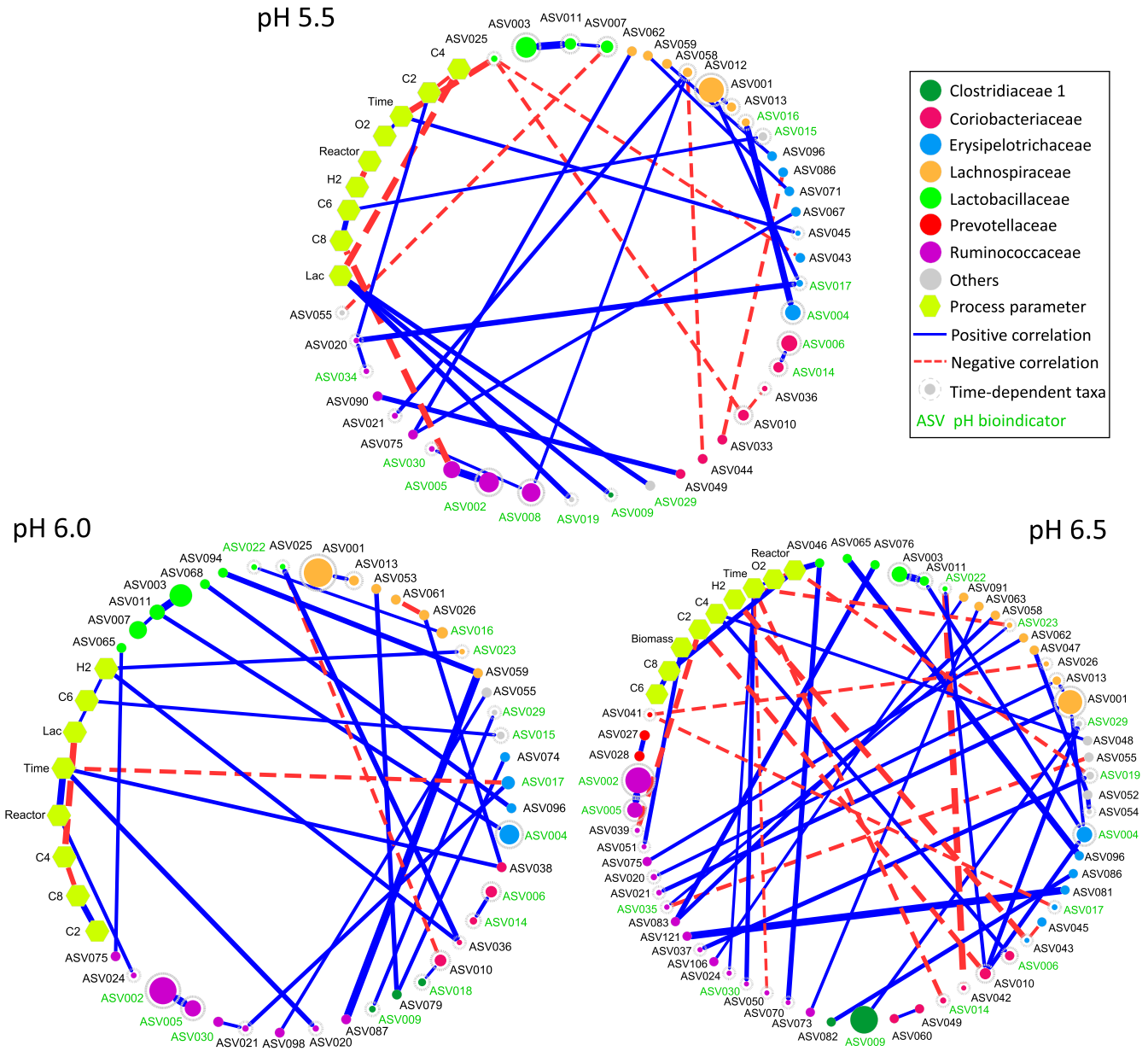


Figure 5. Co-occurrence networks for the three individual pH levels. Edges indicate a coefficient >0.5 for positive correlations and < −0.5 for negative correlations. Edge thickness reflects the strength of the correlation. The size of each ASV node is proportional to the mean relative abundance over the corresponding pH level. ASV nodes are colored and grouped by family. ASV nodes with gray dashed borders are those time-dependent taxa of each individual pH level, whose abundance can be predicted based on the previous microbial community composition. pH bioindicators identified by random forest classification are shown with green letters. "Others" include the ASVs belonging to families *Eubacteriaceae* (ASV015), *Actinomycetaceae* (ASV019), *Clostridiales incertae sedis* XI (ASV029), *Microbacteriaceae* (ASV048), *Veillonellaceae* (ASV052, ASV054), and *Nocardiaceae* (ASV055). Lac, lactate concentration; C2, acetate yield; C4, *n*-butyrate yield; C6, *n*-caproate yield; C8, *n*-caprylate yield.

identification of bioindicators based on microbial community data is a key application of machine learning predictive models.[17] By using RF classification, ASV008 and ASV009 were denoted as pH bioindicators that were most relevant to community dynamics caused by pH increase. Aitchison PCA clustering further highlighted the role of these most influential taxa in driving the community dynamics. By fitting LME models to their relative abundances, we showed that both pH and time significantly affected their dynamics, and pH had a much stronger impact than time. Based on the statistically robust results of Aitchison PCA clustering coupled with LME models and RF classification, a clear conclusion can be drawn: mildly acidic pH values (lower than 6.0) are favorable for *Clostridium* IV while the more neutral pH 6.5 is suitable for *Clostridium sensu stricto*. As described in our previous study, based on similar phylogeny, we linked these ASV bioindicators to the MAGs recovered from the enriched community that served as inoculum for the present study.[17] All five MAGs of *Clostridium* IV and *Clostridium sensu stricto* harbor the genetic potential for CE[17] (Table S9). For *Clostridium* IV ASV008, its corresponding MAGs have 78% average nucleotide identity (ANI) to the lactate-based chain elongator *Ruminococcaceae* bacterium CPB6, which belongs to the family *Acutalibacteraceae* UBA4871 according to the Genome Taxonomy Database.[43] Strain CPB6

https://doi.org/10.1021/acs.est.2c09573
Environ. Sci. Technol. 2023, 57, 18350−18361

was described to prefer mildly acidic pH (5.5−6.0) and to suffer from low growth rates and long lag phases at pH values above 6.0.[44] For *Clostridium sensu stricto* ASV009, its corresponding MAGs showed 81% ANI to *Clostridium luticellarii*, which has a pH optimum of 6.5[45] and CE capability.[46−49] Functional annotation revealed that all genes necessary for lactate oxidation and reverse β-oxidation are present, i.e., these MAGs represent key players of lactate-based CE in our reactor microbiomes.[17] The corresponding ASV008 and ASV009 were identified as time-dependent taxa that are key to understand the community assembly and can be used to characterize the temporal trajectories of the communities. The pH preferences of *Clostridium* IV ASV008 and *Clostridium sensu stricto* ASV009 tied together with concepts in niche theory suggest that different CE bacteria thrive within a defined range of pH values, and outside this range, they are outcompeted by other, better adapted CE species.[50] In our system, the shift of the dominating CE key players happened during the period of pH 6.0 and was completed upon pH 6.5. Due to the distinct growth optima of different populations, alteration of pH is an important tool to shape and control CE reactor microbiomes, in particular when competing reactions, e.g., the consumption of lactate by propionate fermentation, need to be controlled.[15]

pH Value as a Key Determinant of Microbial Community Assembly. Regular and temporally dense sampling with replicates is crucial to capture compositional patterns of communities inferred from time-series data.[7,23] Microbial interaction is one of the main factors affecting such time-dependent patterns. Given that pH had a much stronger association with community assembly than time, we conclude that pH was the main driver modulating microbial interactions. Our former studies indicated that lactate-based CE driven by *Olsenella* is an essential feature when maintaining the pH at 5.5.[14,17] Along with increasing pH, lactic acid bacteria of the genus *Olsenella* cooperating with the chain elongator *Clostridium* IV were replaced by lactic acid bacteria of the genus *Lactobacillus*. Both genera are xylose-fermenting lactate producers according to the functional annotation of their MAGs (Table S9). An enriched community dominated by CE species and *Lactobacillus* was reported in a recent study.[51] Lambrecht et al. suggested inherent benefits of *in situ* lactate formation in CE.[16] The shift in the mutualistic relationship between lactate producers and lactate-consuming chain elongators along the pH gradient revealed the plasticity of the CE microbiota food web. While it is tempting to draw conclusions from observed co-occurrence patterns to further elucidate functional interactions within this food web, care must be taken as species co-occurrence does not necessarily indicate direct metabolic interactions. For example, the co-occurrence of phylogenetically close species may simply indicate their overlapping metabolic niches,[52] such as the appearance of *Lactobacillus* ASV003 and ASV011, *Syntrophococcus* ASV001 and ASV013, and *Clostridium* IV ASV002 and ASV005 at all pH levels. With the increased number of microbial interactions and increasing interaction intensity strongly coupled to the taxa at higher pH, the factor pH shaping the community assembly was revealed by considering the growth and interactions of community members in such long-term closed systems.

Besides, other effects of pH shifts cannot be ignored. At higher pH, the concentrations of protonated carboxylic acids are lower, which are known growth inhibitors of bacteria including CE community members.[11,53−56] The longer the chain length, the more hydrophobic and consequently more toxic the acids are as they can disrupt cell membrane integrity.[56] However, the energy gain for CE bacteria is higher with more CE cycles, i.e., longer-chain products. Notably, both bioreactors showed a transient increase of C8 production after increasing the pH from 5.5 to 6.0. This might be due to the fact that C8 becomes less toxic at higher pH since a greater share is dissociated, facilitating more CE cycles that lead to C8 formation. Thereafter, C8 production dropped to the previous level, which might be explained by the community shifts caused by the pH increase. There are different terminal enzymes catalyzing the reverse β-oxidation and different enzyme complexes involved in energy conservation in CE bacteria, which might have energetic implications for the resulting CE products. Moreover, chain elongation with lactate becomes more exergonic under more acidic pH conditions.[15] For C6 production, thermodynamic analysis suggests that decreasing the pH by one unit releases 3.9 kJ more Gibbs free energy per mole of lactate.

Community Changes Do Not Necessarily Affect Community Functioning. We assumed that an increase in pH would induce shifts in the community assembly and consequently community functioning. However, unlike in a complex, open CE system,[16] increasing pH had no substantial effects on CE community functioning, i.e., changes in community composition did not necessarily lead to improved carboxylate production during long-term reactor operation. This agrees with the rare associations between ASVs and process parameters in the networks. Without introducing new microorganisms by inoculation, the emergence of rare species indicated high functional redundancy despite the reactors being operated as long-term closed systems. According to the storage effect, rare species can germinate and become dominant under proper conditions.[57,58] In this study, the increase in richness can be explained by an abundance shift of some taxa from undetectable to abundant (e.g., *Actinomyces* and *Prevotella* in Figure S6e), reflecting the strong inhibition effects of lower pH on these taxa. Although we operated the reactors in a quasi-continuous mode, which theoretically leads to the washout of organisms that do not grow fast enough, we observed biofilms that unintentionally formed at the glass vessels and could provide a niche for maintaining such rare populations even under conditions that do not favor their growth. The reactor performance returned to the previous state after the fluctuation in carboxylate production along pH gradients, which might be due to overlapping metabolic niches with coexisting rare species that could increase functional resilience to environmental disturbances. As mentioned above, the pH shift caused a dramatic but transient increase of C8 yield. How to exploit such disturbance effects for process control needs to be investigated systematically. Keeping functional redundancy in mixed culture processes might be important for biotechnological applications, because parallel pathways of substrate conversion are essential to guarantee the functional stability during perturbation.[9,11,59] With regard to the practical implications for mixed-culture production processes within the carboxylate platform, our results delineate fundamental differences between long-term enriched microbiomes selected for the production of C6/C8 and engineered consortia assembled from single species covering all metabolic traits needed for that function. The latter might perform better under stable conditions, whereas naturally selected consortia keeping rare species are more robust under fluctuating conditions and resilient toward perturbations due to their functional redundancy. Efficient microbial resource management is of paramount importance for implementing

https://doi.org/10.1021/acs.est.2c09573
Environ. Sci. Technol. 2023, 57, 18350−18361

mixed-culture bioprocesses for the carboxylate platform at industrial scale, as the valorization of organic residues and feedstocks of fluctuating quality requires knowledge-based community engineering.

■ ASSOCIATED CONTENT

Data Availability Statement

Amplicon sequencing data (ERR4450775−ERR4450910) have been deposited to the ENA database under study no. PRJEB39808. The MAG sequences used in this study are publicly available in ENA under the accession nos. GCA_903789645, GCA_903789675, GCA_903789585, GCA_903789565, GCA_903789665, GCA_903789475, GCA_903789455, GCA_903789485, GCA_903789575, GCA_903789695, and GCA_903789705.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.2c09573.

Figures of α-diversity rarefaction curves, kinetics of xylan consumption in the bioreactors within 24 hours, gas production of the bioreactors, biomass production of the bioreactors, longitudinal changes in diversity and evenness of order two of bioreactor communities, microbial community composition profiles of the bioreactors, dissimilarities in bacterial community composition (β-diversity) of the two bioreactors A and B, variation partitioning analysis showing the relative importance of pH and time on microbial community variations, nested cross-validation of random forest classification in the prediction of pH levels for each sample, core time-dependent taxa of individual pH levels, and co-occurrence network for the entire period of reactor operation and tables of linear mixed-effects model results for diversity, evenness, and richness of order one, linear mixed-effects model results for the relative abundance of *Clostridium* IV ASV008 and *Clostridium sensu stricto* ASV009 at the different pH levels, linear mixed-effects model results for microbial community composition that is represented by the PC1 from the Aitchison and Bray-Curtis distance-based principal component analysis, summary statistics of networks, and metagenome-assembled genomes with the same taxonomy as ASVs (PDF)

Process parameter data sets used for modeling and ASV table (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

Sabine Kleinsteuber − *Department of Environmental Microbiology, Helmholtz Centre for Environmental Research − UFZ, 04318 Leipzig, Germany;* orcid.org/0000-0002-8643-340X; Phone: +49 341 235-1325; Email: sabine.kleinsteuber@ufz.de

Authors

Bin Liu − *Department of Environmental Microbiology, Helmholtz Centre for Environmental Research − UFZ, 04318 Leipzig, Germany; KU Leuven, Department of Microbiology, Immunology and Transplantation, Rega Institute for Medical Research, Laboratory of Molecular Bacteriology, BE-3000 Leuven, Belgium*

Heike Sträuber − *Department of Environmental Microbiology, Helmholtz Centre for Environmental Research − UFZ, 04318 Leipzig, Germany*

Florian Centler − *Department of Environmental Microbiology, Helmholtz Centre for Environmental Research − UFZ, 04318 Leipzig, Germany; School of Life Sciences, University of Siegen, 57076 Siegen, Germany*

Hauke Harms − *Department of Environmental Microbiology, Helmholtz Centre for Environmental Research − UFZ, 04318 Leipzig, Germany*

Ulisses Nunes da Rocha − *Department of Environmental Microbiology, Helmholtz Centre for Environmental Research − UFZ, 04318 Leipzig, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.2c09573

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Ute Lohse for skilled technical assistance in molecular analyses and the colleagues from DBFZ Deutsches Biomasseforschungszentrum gGmbH for their technical support in analyses of abiotic parameters. The authors thank Wanwan Liang from the Centre of Geospatial Analytics at North Carolina State University and Liat Shenhav from Department of Computer Science at University of California Los Angeles for their help with statistics. B.L. was funded by the China Scholarship Council (# 201606350010). F.C., H.S., and S.K. were supported by the BMBF − German Federal Ministry of Education and Research (# 031B0389B, # 01DQ17016, and # 031A317) and the Helmholtz Association (Program Renewable Energies). U.N.d.R. was financed by the Helmholtz Young Investigator grant VH-NG-1248 Micro "Big Data".

■ REFERENCES

(1) Goldford, J. E.; Lu, N.; Bajić, D.; Estrela, S.; Tikhonov, M.; Sanchez-Gorostiaga, A.; Segrè, D.; Mehta, P.; Sanchez, A. Emergent simplicity in microbial community assembly. *Science* **2018**, *361*, 469−474.

(2) Nemergut, D. R.; Schmidt, S. K.; Fukami, T.; O'Neill, S. P.; Bilinski, T. M.; Stanish, L. F.; Knelman, J. E.; Darcy, J. L.; Lynch, R. C.; Wickey, P.; Ferrenberg, S. Patterns and processes of microbial community assembly. *Microbiol. Mol. Biol. Rev.* **2013**, *77* (3), 342−356.

(3) Antwis, R. E.; Griffiths, S. M.; Harrison, X. A.; Aranega-Bou, P.; Arce, A.; Bettridge, A. S.; Brailsford, F. L.; de Menezes, A.; Devaynes, A.; Forbes, K. M.; Fry, E. L.; Goodhead, I.; Haskell, E.; Heys, C.; James, C.; Johnston, S. R.; Lewis, G. R.; Lewis, Z.; Macey, M. C.; McCarthy, A.; McDonald, J. E.; Mejia-Florez, N. L.; O'Brien, D.; Orland, C.; Pautasso, M.; Reid, W. D. K.; Robinson, H. A.; Wilson, K.; Sutherland, W. J. Fifty important research questions in microbial ecology. *FEMS Microbiol. Ecol.* **2017**, *93*, fix044.

(4) Tripathi, B. M.; Stegen, J. C.; Kim, M.; Dong, K.; Adams, J. M.; Lee, Y. K. Soil pH mediates the balance between stochastic and deterministic assembly of bacteria. *ISME J.* **2018**, *12* (4), 1072−1083.

(5) Fargione, J.; Brown, C. S.; Tilman, D. Community assembly and invasion: an experimental test of neutral versus niche processes. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 8916−8920.

(6) Chesson, P. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* **2000**, *31*, 343−366.

(7) Faust, K.; Bauchinger, F.; Laroche, B.; de Buyl, S.; Lahti, L.; Washburne, A. D.; Gonze, D.; Widder, S. Signatures of ecological processes in microbial community time series. *Microbiome* **2018**, *6* (1), 120.

(8) Wu, L.; Ning, D.; Zhang, B.; Li, Y.; Zhang, P.; Shan, X.; Zhang, Q.; Brown, M.; Li, Z.; Van Nostrand, J. D.; Ling, F.; Xiao, N.; Zhang, Y.; Vierheilig, J.; Wells, G. F.; Yang, Y.; Deng, Y.; Tu, Q.; Wang, A.; Zhang, T.; He, Z.; Keller, J.; Nielsen, P. H.; Alvarez, P. J. J.; Criddle, C. S.; Wagner, M.; Tiedje, J. M.; He, Q.; Curtis, T. P.; Stahl, D. A.; Alvarez-Cohen, L.; Rittmann, B. E.; Wen, X.; Zhou, J. Global diversity and biogeography of bacterial communities in wastewater treatment plants. *Nat. Microbiol.* **2019**, *4*, 1183−1195.
(9) Werner, J. J.; Knights, D.; Garcia, M. L.; Scalfone, N. B.; Smith, S.; Yarasheski, K.; Cummings, T. A.; Beers, A. R.; Knight, R.; Angenent, L. T. Bacterial community structures are unique and resilient in full-scale bioenergy systems. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (10), 4158−4163.
(10) Kucek, L. A.; Nguyen, M.; Angenent, L. T. Conversion of L-lactate into n-caproate by a continuously fed reactor microbiome. *Water Res.* **2016**, *93*, 163−171.
(11) Agler, M. T.; Werner, J. J.; Iten, L. B.; Dekker, A.; Cotta, M. A.; Dien, B. S.; Angenent, L. T. Shaping reactor microbiomes to produce the fuel precursor n-butyrate from pretreated cellulosic hydrolysates. *Environ. Sci. Technol.* **2012**, *46* (18), 10229−10238.
(12) Angenent, L. T.; Richter, H.; Buckel, W.; Spirito, C. M.; Steinbusch, K. J. J.; Plugge, C. M.; Strik, D. P. B. T. B.; Grootscholten, T. I. M.; Buisman, C. J. N.; Hamelers, H. V. M. Chain elongation with reactor microbiomes: Open-culture biotechnology to produce biochemicals. *Environ. Sci. Technol.* **2016**, *50* (6), 2796−2810.
(13) Agler, M. T.; Wrenn, B. A.; Zinder, S. H.; Angenent, L. T. Waste to bioproduct conversion with undefined mixed cultures: The carboxylate platform. *Trends Biotechnol.* **2011**, *29* (2), 70−78.
(14) Liu, B.; Kleinsteuber, S.; Centler, F.; Harms, H.; Sträuber, H. Competition between butyrate fermenters and chain-elongating bacteria limits the efficiency of medium-chain carboxylate production. *Front. Microbiol.* **2020**, *11*, 336.
(15) Candry, P.; Radić, L.; Favere, J.; Carvajal-Arroyo, J. M.; Rabaey, K.; Ganigué, R. Mildly acidic pH selects for chain elongation to caproic acid over alternative pathways during lactic acid fermentation. *Water Res.* **2020**, *186*, 116396.
(16) Lambrecht, J.; Cichocki, N.; Schattenberg, F.; Kleinsteuber, S.; Harms, H.; Müller, S.; Sträuber, H. Key sub-community dynamics of medium-chain carboxylate production. *Microb. Cell Fact.* **2019**, *18* (1), 92.
(17) Liu, B.; Sträuber, H.; Saraiva, J.; Harms, H.; Silva, S. G.; Kasmanas, J. C.; Kleinsteuber, S.; Nunes da Rocha, U. Machine learning-assisted identification of bioindicators predicts medium-chain carboxylate production performance of an anaerobic mixed culture. *Microbiome* **2022**, *10*, 48.
(18) Lawson, C. E.; Harcombe, W. R.; Hatzenpichler, R. Common principles and best practices for engineering microbiomes. *Nat. Rev. Microbiol.* **2019**, *17*, 725−741.
(19) Gloor, G. B.; Macklaim, J. M.; Pawlowsky-Glahn, V.; Egozcue, J. J. Microbiome datasets are compositional: And this is not optional. *Front. Microbiol.* **2017**, *8*, 2224.
(20) Martino, C.; Morton, J. T.; Marotz, C. A.; Thompson, L. R.; Tripathi, A.; Knight, R.; Zengler, K. A novel sparse compositional technique reveals microbial perturbations. *mSystems* **2019**, *4* (1), No. e00016-19.
(21) Anderson, M. J. A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* **2001**, *26*, 32−46.
(22) Tackmann, J.; Rodrigues, J. F. M.; von Mering, C. Rapid inference of direct interactions in large-scale ecological networks from heterogeneous microbial sequencing data. *Cell Syst.* **2019**, *9* (3), 286−296.
(23) Faust, K.; Lahti, L.; Gonze, D.; de Vos, W. M.; Raes, J. Metagenomics meets time series analysis: Unraveling microbial community dynamics. *Curr. Opin. Microbiol.* **2015**, *25* (May), 56−66.
(24) Ryo, M.; Aguilar-Trigueros, C. A.; Pinek, L.; Muller, L. A. H.; Rillig, M. C. Basic principles of temporal dynamics. *Trends Ecol. Evol.* **2019**, *34* (8), 723−733.
(25) Shenhav, L.; Furman, O.; Briscoe, L.; Thompson, M.; Silverman, J. D.; Mizrahi, I.; Halperin, E. Modeling the temporal dynamics of the gut microbial community in adults and infants. *PLoS Comput. Biol.* **2019**, *15*, No. e1006960.
(26) Bokulich, N. A.; Dillon, M. R.; Zhang, Y.; Rideout, J. R.; Bolyen, E.; Li, H.; Albert, P. S.; Caporaso, J. G. Q2-Longitudinal: Longitudinal and paired-sample analyses of microbiome data. *mSystems* **2018**, *3* (6), No. e00219-18.
(27) Sträuber, H.; Bühligen, F.; Kleinsteuber, S.; Dittrich-Zechendorf, M. Carboxylic acid production from ensiled crops in anaerobic solid-state fermentation - Trace elements as pH controlling agents support microbial chain elongation with lactic acid. *Eng. Life Sci.* **2018**, *18* (7), 447−458.
(28) Urban, C.; Xu, J.; Sträuber, H.; dos Santos Dantas, T. R.; Mühlenberg, J.; Härtig, C.; Angenent, L. T.; Harnisch, F. Production of drop-in fuel from biomass by combined microbial and electrochemical conversions. *Energy Environ. Sci.* **2017**, *10*, 2231−2244.
(29) Lucas, R.; Kuchenbuch, A.; Fetzer, I.; Harms, H.; Kleinsteuber, S. Long-term monitoring reveals stable and remarkably similar microbial communities in parallel full-scale biogas reactors digesting energy crops. *FEMS Microbiol. Ecol.* **2015**, *91* (3), fiv004.
(30) Klindworth, A.; Pruesse, E.; Schweer, T.; Peplies, J.; Quast, C.; Horn, M.; Glöckner, F. O. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **2013**, *41*, No. e1.
(31) Bolyen, E.; Rideout, J. R.; Dillon, M. R.; Bokulich, N. A.; Chase, J.; Cope, E. K.; Gorlick, K.; Keefe, C. R.; Keim, P.; Kreps, J.; Naimey, A. T.; Pearson, T.; Shiffer, A.; Caporaso, J. G.; Abnet, C. C.; Loftfield, E.; Sinha, R.; Vogtmann, E.; Wan, Y.; Al-Ghalith, G. A.; Alexander, H.; Brown, C. T.; Alm, E. J.; Duvallet, C.; Alm, E. J.; Arumugam, M.; Brejnrod, A.; Rasmussen, L. B.; Asnicar, F.; Segata, N.; Bai, Y.; Liu, Y. X.; Bai, Y.; Liu, Y. X.; Bisanz, J. E.; Bittinger, K.; Brislawn, C. J.; Callahan, B. J.; Callahan, B. J.; Caraballo-Rodríguez, A. M.; Da Silva, R.; Dorrestein, P. C.; Ernst, M.; Gauglitz, J. M.; Jarmusch, A. K.; Kang, K. B.; Koester, I.; Melnik, A. V.; Nothias, L. F.; Petras, D.; Tripathi, A.; Wang, M.; Caporaso, J. G.; Diener, C.; Gibbons, S. M.; Douglas, G. M.; Durall, D. M.; Edwardson, C. F.; Estaki, M.; Fouquier, J.; Shaffer, M.; Fouquier, J.; Lozupone, C.; Pruesse, E.; Gibbons, S. M.; Gibson, D. L.; Gonzalez, A.; Holste, H.; Marotz, C.; McDonald, D.; Morton, J. T.; Navas-Molina, J. A.; Song, S. J.; Vázquez-Baeza, Y.; Xu, Z. Z.; Zhu, Q.; Knight, R.; Guo, J.; Hillmann, B.; Knights, D.; Holmes, S.; Holste, H.; Huttenhower, C.; McIver, L. J.; Huttenhower, C.; Huttley, G. A.; Kaehler, B. D.; Janssen, S.; Jiang, L.; Kang, K. Bin; Kelley, S. T.; Kosciolek, T.; Langille, M. G. I.; Lee, J.; Ley, R.; Walters, W.; Maher, M.; Navas-Molina, J. A.; Martin, B. D.; Metcalf, J. L. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **2019**, *37* (8), 852−857.
(32) Callahan, B. J.; McMurdie, P. J.; Rosen, M. J.; Han, A. W.; Johnson, A. J. A.; Holmes, S. P. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, *13* (7), 581−583.
(33) McIlroy, S. J.; Kirkegaard, R. H.; McIlroy, B.; Nierychlo, M.; Kristensen, J. M.; Karst, S. M.; Albertsen, M.; Nielsen, P. H. MiDAS 2.0: An ecosystem-specific taxonomy and online database for the organisms of wastewater treatment systems expanded for anaerobic digester groups. *Database* **2017**, *2017*, bax016.
(34) Wang, Q.; Garrity, G. M.; Tiedje, J. M.; Cole, J. R. Naıve Bayesian classifier for rapid assignment of rRNA sequences. *Appl. Environ. Microbiol.* **2007**, *73*, 5261−5267.
(35) Sanders, H. L. Marine benthic diversity: A comparative study. *Am. Nat.* **1968**, *102* (925), 243.
(36) Willis, A. D. Rarefaction, alpha diversity, and statistics. *Front. Microbiol.* **2019**, *10*, 2407.
(37) Lucas, R.; Groeneveld, J.; Harms, H.; Johst, K.; Frank, K.; Kleinsteuber, S. A critical evaluation of ecological indices for the comparative analysis of microbial communities based on molecular datasets. *FEMS Microbiol. Ecol.* **2017**, *93* (1), fiw209.
(38) Bray, J. R.; Curtis, J. T. An ordination of the upland forest communities of Southern Wisconsin. *Ecol. Monogr.* **1957**, *27* (4), 325−349.

(39) Fedarko, M. W.; Martino, C.; Morton, J. T.; González, A.; Rahman, G.; Marotz, C. A.; Minich, J. J.; Allen, E. E.; Knight, R. Visualizing 'omic feature rankings and log-ratios using Qurro. *NAR Genomics Bioinforma.* **2020**, *2* (2), lqaa023.
(40) Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **1995**, *57* (1), 289−300.
(41) Bokulich, N.; Dillon, M.; Bolyen, E.; Kaehler, B.; Huttley, G.; Caporaso, J. q2-sample-classifier: machine-learning tools for microbiome classification and regression. *J. Open Source Softw.* **2018**, *3* (30), 934.
(42) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498−2504.
(43) Parks, D. H.; Chuvochina, M.; Waite, D. W.; Rinke, C.; Skarshewski, A.; Chaumeil, P. A.; Hugenholtz, P. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **2018**, *36*, 996.
(44) Zhu, X.; Zhou, Y.; Wang, Y.; Wu, T.; Li, X.; Li, D.; Tao, Y. Production of high-concentration *n*-caproic acid from lactate through fermentation using a newly isolated *Ruminococcaceae* bacterium CPB6. *Biotechnol. Biofuels* **2017**, *10* (1), 102.
(45) Wang, Q.; Wang, C. D.; Li, C. H.; Li, J. G.; Chen, Q.; Li, Y. Z. *Clostridium luticellarii* sp. nov., isolated from a mud cellar used for producing strong aromatic liquors. *Int. J. Syst. Evol. Microbiol.* **2015**, *65* (12), 4730−4733.
(46) Van Brabant, P. *Understanding bio-isomerisation during methanol fermentation*. Master Thesis, Universiteit Gent, 2019.
(47) de Leeuw, K. D.; de Smit, S. M.; van Oossanen, S.; Moerland, M. J.; Buisman, C. J. N.; Strik, D. P. B. T. B. Methanol-based chain elongation with acetate to n-butyrate and isobutyrate at varying selectivities dependent on pH. *ACS Sustain. Chem. Eng.* **2020**, *8* (22), 8184−8194.
(48) de Smit, S. M.; de Leeuw, K. D.; Buisman, C. J. N.; Strik, D. P. B. T. B. Continuous n-valerate formation from propionate and methanol in an anaerobic chain elongation open-culture bioreactor. *Biotechnol. Biofuels* **2019**, *12* (1), 132.
(49) Huang, S.; Kleerebezem, R.; Rabaey, K.; Ganigué, R. Open microbiome dominated by *Clostridium* and *Eubacterium* converts methanol into i-butyrate and n-butyrate. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 5119−5131.
(50) Morton, J. T.; Sanders, J.; Quinn, R. A.; McDonald, D.; Gonzalez, A.; Vázquez-Baeza, Y.; Navas-Molina, J. A.; Song, S. J.; Metcalf, J. L.; Hyde, E. R.; Lladser, M.; Dorrestein, P. C.; Knight, R. Balance trees reveal microbial niche differentiation. *mSystems* **2017**, *2* (1), No. e00162-16.
(51) Contreras-Dávila, C. A.; Carrión, V. J.; Vonk, V. R.; Buisman, C. N. J.; Strik, D. P. B. T. B. Consecutive lactate formation and chain elongation to reduce exogenous chemicals input in repeated-batch food waste fermentation. *Water Res.* **2020**, *169*, 115215.
(52) Zelezniak, A.; Andrejev, S.; Ponomarova, O.; Mende, D. R.; Bork, P.; Patil, K. R. Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (20), 6449−6454.
(53) Andersen, S. J.; de Groof, V.; Khor, W. C.; Roume, H.; Props, R.; Coma, M.; Rabaey, K. A *Clostridium* Group IV species dominates and suppresses a mixed culture fermentation by tolerance to medium chain fatty acids products. *Front. Bioeng. Biotechnol.* **2017**, *5*, 8.
(54) Ge, S.; Usack, J. G.; Spirito, C. M.; Angenent, L. T. Long-term n-caproic acid production from yeast-fermentation beer in an anaerobic bioreactor with continuous product extraction. *Environ. Sci. Technol.* **2015**, *49* (13), 8012−8021.
(55) Royce, L. A.; Liu, P.; Stebbins, M. J.; Hanson, B. C.; Jarboe, L. R. The damaging effects of short chain fatty acids on *Escherichia coli* membranes. *Appl. Microbiol. Biotechnol.* **2013**, *97* (18), 8317−8327.
(56) Jarboe, L. R.; Royce, L. A.; Liu, P. Understanding biocatalyst inhibition by carboxylic acids. *Front. Microbiol.* **2013**, *4*, 272.
(57) Chesson, P. L.; Warner, R. R. Environmental variability promotes coexistence in lottery competitive systems. *Am. Nat.* **1981**, *117*, 923−943.
(58) Lennon, J. T.; Jones, S. E. Microbial seed banks: The ecological and evolutionary implications of dormancy. *Nat. Rev. Microbiol.* **2011**, *9* (2), 119−130.
(59) Hashsham, S. A.; Fernandez, A. S.; Dollhopf, S. L.; Dazzo, F. B.; Hickey, R. F.; Tiedje, J. M.; Criddle, C. S. Parallel processing of substrate correlates with greater functional stability in methanogenic bioreactor communities perturbed by glucose. *Appl. Environ. Microbiol.* **2000**, *66* (9), 4050−4057.

https://doi.org/10.1021/acs.est.2c09573
Environ. Sci. Technol. 2023, 57, 18350−18361