Data in Brief 57 (2024) 111095

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Abstract and composition data of microbiomes in agricultural biogas plants of Lower Saxony, Germany, with variation in organic substrates, process parameters and nutrients

Sascha M.B. Krause[a,b], Rui Wang[b], Anja B. Dohrmann[a,c], Meike Walz[d], Achim Loewen[d], Christoph C. Tebbe[a,*]

[a] *Thünen Institute of Biodiversity, Federal Research Institute for Rural Areas, Forestry and Fisheries, Bundesallee 65, Braunschweig 38116, Germany*
[b] *School of Ecological and Environmental Sciences, East China Normal University, Dongchuan Road 500, Shanghai 200241, China*
[c] *Federal Institute for Geosciences and Natural Resources (BGR), Stilleweg 2, 30655 Hannover, Germany*
[d] *Faculty of Resource Management, University of Applied Sciences and Arts (HAWK), Rudolf-Diesel-Str. 12, 37075 Göttingen, Germany*

ARTICLE INFO

Article history:
Received 28 May 2024
Revised 23 October 2024
Accepted 25 October 2024
Available online 4 November 2024

Dataset link: Dataset: Abundance and composition of microbiomes in agricultural biogas plants in Lower Saxony, Germany, with variation in organic substrates, process parameters and nutrients. (Original data)
Dataset link: F2_DADA2_ASV_data.xlsx (Original data)

ABSTRACT

This article presents high-throughput DNA sequencing, quantitative PCR data of microbial communities, and process parameters as recovered from eight biogas plants (BPs) located in Lower Saxony, Germany. Samples were collected from both the main (MD) and secondary digesters (SD). Additionally, for 4 BPs, samples were also obtained from the residue digester storage (RDS). Different BPs employed various types of substrates originating from cattle manure, chicken manure, pig manure, or renewable resources. Information on physico-chemical process parameters and concentrations of macro- and micro-nutrients in the BPs is provided. Total DNA from all samples were extracted using a phenol-chloroform-based method. To determine the abundance of bacteria and archaea, their 16S rRNA genes were quantified by real-time PCR

* Corresponding author.
E-mail address: christoph.tebbe@thuenen.de (C.C. Tebbe).
Social media: @ChristophTebbe (C.C. Tebbe)

https://doi.org/10.1016/j.dib.2024.111095

Keywords:
Agricultural biogas plants
Anaerobic microbiome
Microbial diversity
PCR 16S rRNA gene analyses
DNA sequencing
Quantitative PCR
Organic waste treatment
Bioenergy

(qPCR), and to characterize their community composition, paired-end DNA sequence reads were generated from PCR amplicons with Illumina MiSeq. All statistical analyses were performed in R to explore the microbial diversity, abundance, and community structure among different BPs and digesters (MD, SD, RDS). The presence and distribution of the major bacterial and archaeal phyla indicated for each BP unique and diverse microbial communities with typically higher bacterial than archaeal abundances.

Specifications Table

Subject	Environmental Sciences; Agricultural Sciences
Specific subject area	Molecular Ecology, microbial composition and diversity, anaerobic digester process parameters
Data format	Raw fastq files, Analysed, Filtered
Type of data	16S amplicon metagenomic reads, 3 Tables, 4 Figures
Data collection	Samples were first subjected to nucleic acid extraction, followed by 16S rRNA qPCR and gene analysis using MiSeq technology. The raw data was processed through the DADA2 pipeline. Data was normalized by a Bayesian-multiplicative replacement, count zero multiplicative (CZM) method and centered log-ratio transformation.
Data source location	All biogas plants are located in Lower Saxony, Federal State of Germany
Data accessibility	A. Repository name for qPCR and analyses of the microbial community compositions: OpenAgrarData identification number: DOI 10.3220/DATA20240919153413-0 Direct URL to data: https://doi.org/10.3220/DATA20240919153413-0 B. Repository name for DNA sequences: European Nucleotide ArchiveData identification number: PRJEB75859 https://www.ebi.ac.uk/ena/browser/view/PRJEB75859
Related research article	A.B. Dohrmann, M. Walz, Loewen A, C. Tebbe, *Clostridium* cluster I and their pathogenic members in a full-scale operating biogas plant, Appl. Microbiol. Biotechnol. 99 (2015) 3585-3598. https://doi.org/10.1007/s00253-014-6261-y

1. Value of the Data

- We report physicochemical process parameters and concentrations of macro- and micro-nutrients for eight biogas plants (BPs), including samples from the main (MD) and secondary digesters (SD), as well as residue digester storage (RDS). This data allows researchers to assess the environmental conditions that influence biogas production efficiency and microbial activity across different plants.
- The sequencing and qPCR data provides comprehensive insights into the abundance and diversity of archaea and bacteria, the key microbial groups driving the biogas formation process. By comparing microbial communities across various substrates and plant locations, this dataset enables researchers to explore how different environmental and operational conditions impact microbial community structure and function.
- Using commonly employed primers 515F (5′-GTGYCAGCMGCCGCGGTA-3′) and 909R (5′-CCCCGYCAATTCMTTTRAGT-3′) ensures that this dataset can be compared to other similar studies, facilitating broader meta-analyses and cross-study comparisons.
- By combining sequencing data with quantitative PCR, this dataset offers a valuable resource for studying variations in the biogas formation process. It enables researchers to investigate

microbial population dynamics and use these insights to develop control strategies to optimize biogas production efficiency.

2. Background

The anaerobic fermentation of organic substances and waste to produce methane in biogas reactors has become an important energy source of considerable value for farms in Lower Saxony, Germany. The actual fermentation process involves an anaerobic microbiome that contributes to sequential metabolic steps, i.e., the decomposition of cellulolytic and other polymeric substances to monomers, the fermentation of such monomers by primary fermenters to acetate, and finally methanogenesis. Each of these steps is carried out by specific taxa, usually including a high proportion of Clostridia and methanogenic archaea, but the variation of these taxonomic groups in response to different organic substrates entering the process and specific process conditions at farm scale is poorly understood. The data presented here a give snapshot overview of the variability of the microbiome compositions as it was detected from biogas reactors at eight different farms in Lower Saxony. Compositional comparisons are made and co-occurring physiochemical conditions of the substrates are reported. The data can be useful for future meta-analyses showing results from other agricultural biogas facility (Tables 1-3).

Table 1
Overview of the sampled BPs and their substrates. Main digester (MD) and secondary digester (SD) were sampled at the BPs P1, P3, P4 and P5. In the P2 BP, only the MD was sampled. In the P6-P8 BP, in addition to the MD and SD, the residue digester storage (RDS) was also sampled.

Label	Substrate composition
Renewable resources	
P1b	Maize silage, whole plant cereal silage
Cattle manure	
P1a	Maize silage, whole plant cereal silage, cattle manure
P6	Maize silage, grass silage, cattle manure
P8	Maize silage, Leftover feed, cattle manure
Pig manure	
P5	Maize silage, cattle and pig manure, pig faeces
P7	Maize silage, pig manure, dry chicken manure
Chicken manure	
P2	Maize silage, sugar beet, beef manure, dry chicken manure
P3	Maize silage, cattle and pig faeces, pig manure, dry chicken manure
P4	Maize silage, whole plant cereal silage, sugar beet, pig and cattle manure, dry chicken manure

3. Data Description

The Illumina MiSeq platform was used to generate paired-end sequence reads from PCR amplicons of the V4–V5 regions of 16S rRNA gene. Two batches of sequencing were performed, with one covering BPs P1, P3, P4, and P5, and the other BPs P2, P6, P7, and P8. A sample was collected for each type of digester, was split into two aliquots, and were separately analysed. From each DNA extract, three libraries were generated. This resulted in a total of six libraries for each sample. The DADA2 pipeline [1] was employed to process the raw 16S rRNA gene sequences and generate amplicon sequence variants (ASV). The filtered ASV table was used for calculating the Shannon diversity, relative abundance analysis, and Aitchison distance based non-metric multidimensional scaling (NMDS). Quantitative PCR data was obtained by using a StepOnePlus Real-Time PCR system (Life Technologies GmbH) using primers and probes described by Yu and colleagues [2] (Figs. 1-4).

Table 2
Recorded process parameters of BPs P1-P8 (see Table 1 for details).

Biogas plant		P1a MD	P1a SD	P1b MD	P1b SD	P2 MD	P3 MD	P3 SD	P4 MD	P4 SD
Sampling		June 5, 2012				Dec 10, 2012	May 14, 2013		Dec 2, 2013	
pH	[-]	7.6	7.6	7.6	7.7	7.9	8.0	8.1	7.7	7.7
Volatile organic acids (VOA)	[mg/L]	1, 917	1, 917	1, 834	2, 083	3, 079	4, 490	3, 494	2, 249	2, 083
Total inorganic carbon (TAC)	[mg/L]	11, 600	11, 600	12, 000	12, 300	16, 250	19, 000	20, 400	11, 300	11, 950
VOA/TAC	[-]	0.17	0.16	0.15	0.17	0.19	0.24	0.17	0.20	0.17
Dry matter (DM)	[%WW]	9.0	7.7	8.9	7.7	10.7	12.3	10.3	9.9	7.9
Organic dry matter	[%DW]	82.0	79.8	81.9	79.5	66.6	n.b	n.b.	66.5	62.3
NH_4-N	[mg/L]	1, 727	1, 925	1, 786	1, 997	2, 834	3, 603	3, 995	1, 609	1, 878
Lactic acid	[mg/L]	92	92	92	99	< 30	289	227	< 30	35
Acetic acid	[mg/L]	43	< 30	52	< 30	32	976	191	105	100
Propionic acid	[mg/L]	< 30	< 30	< 30	< 30	< 30	175	53	< 30	< 30
Isobutyric acid	[mg/L]	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	74
Butyric acid	[mg/L]	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30
Iso-valeric acid	[mg/L]	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30
Valeric acid	[mg/L]	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30
Caproic acid	[mg/L]	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30

Table 2 (continued).

Biogas plant		P5 MD	P5 SD	P5 RDS	P6 MD	P6 SD	P6 RDS	P7 MD	P7 SD	P7 RDS	P8 MD	P8 SD	P8 RDS
Sampling		Feb 3, 2014			July 23, 2014						July 29, 2014		
pH	[-]	7.7	8.1	7.6	7.9	7.9	7.9	7.9	8.0	7.9	7.8	7.8	7.3
Volatile organic acids (VOA)	[mg/L]	3, 328	2, 249	1, 917	1, 419	1, 751	1, 585	2, 913	1, 917	1, 917	2, 166	2, 000	755
Total inorganic carbon (TAC)	[mg/L]	10, 000	13, 000	12, 100	11, 750	11, 350	11, 200	13, 300	14, 450	14, 350	10, 250	11, 400	4, 450
VOA/TAC	[-]	0.33	0.17	0.16	0.12	0.15	0.14	0.22	0.13	0.13	0.21	0.18	0.17
Dry matter (DM)	[%WW]	8.5	6.1	5.0	5.9	6.2	5.7	7.4	6.7	4.9	6.6	5.5	2.7
Organic dry matter	[%DW]	81.8	76.6	75.9	72.4	72.9	73.3	78.0	75.5	71.9	70.56	71.23	73.2
NH_4-N	[mg/L]	1, 546	2, 268	2, 200	1, 913	1, 847	1, 869	2, 668	2, 865	2, 926	1, 603	1, 891	1, 205
Lactic acid	[mg/L]	167	146	119	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30
Acetic acid	[mg/L]	1427	47	66	< 30	< 30	< 30	437	< 30	< 30	190	38	35
Propionic acid	[mg/L]	198	79	44	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30
Isobutyric acid	[mg/L]	< 30	< 30	< 30	< 30	< 30	< 30	48	< 30	< 30	< 30	< 30	< 30
Butyric acid	[mg/L]	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30
Iso-valeric acid	[mg/L]	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30
Valeric acid	[mg/L]	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30
Caproic acid	[mg/L]	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30	< 30

Table 3
Micro- and macro-nutrients of BPs P1-P8 (see Table 1 for details).

Biogas plant		P1a MD	P1a SD	P1b MD	P1b SD	P2 MD	P3 MD	P3 SD	P4 MD	P4 SD
Sampling		June 5, 2012			Dec 10, 2012		May 14, 2013		Dec 2, 2013	
DW	[%WW]	9.0	7.7	8.9	7.7	10.7	12.3	10.3	9.9	7.9
oDW	[%DW]	82.0	79.8	81.9	79.5	66.6	0.0	0.0	66.5	62.3
Calcium	[mg/kg TS]	13, 840	12, 600	11, 850	12, 900	22, 220	30, 200	26, 300	12, 563	13, 965
Potassium	[mg/kg DW]	46, 750	50, 000	44, 300	51, 900	51, 460	41, 100	46, 800	47, 995	56, 369
Magnesium	[mg/kg DW]	6, 020	6, 340	5, 620	6, 750	1, 760	7, 600	7, 900	5, 208	6, 014
Iron	[mg/kg DW]	2, 010	1, 950	1, 670	2, 060	3, 940	2, 080	2, 270	4, 040	4, 407
Sodium	[mg/kg DW]	1, 430	1, 460	1, 140	1, 180	3, 290	3, 100	3, 500	2, 481	2, 753
Manganese	[mg/kg DW]	36	238	203	245	556	255	268	285	318
Copper	[mg/kg DW]	39	38	32	42	38	59	92	32	41
Zinc	[mg/kg DW]	86	254	189	221	428	320	330	127	137
Nickel	[mg/kg DW]	3.49	3.59	3.13	3.68	6.99	3.11	4.35	11.10	13.10
Molybdenum	[mg/kg DW]	2.37	2.97	2.42	2.70	2.55	3.14	3.59	4.10	4.80
Cobalt	[mg/kg DW]	1.69	1.81	1.46	1.86	1.51	1.16	1.22	3.00	3.40
Selenium	[mg/kg DW]	0.00	0.00	0.00	0.00	0.00	0.68	0.78	0.30	0.70
Calcium	[mg/kg WW]	1, 241	975	1, 052	992	2, 371	3, 709	2, 709	1, 244	1, 103
Potassium	[mg/kg WW]	4, 193	3, 870	3, 934	3, 991	5, 491	5, 047	4, 820	4, 752	4, 453
Magnesium	[mg/kg WW]	540	491	499	519	188	933	814	516	475
Iron	[mg/kg WW]	180	151	148	158	420	255	234	400	348
Sodium	[mg/kg WW]	128	113	101	91	351	381	361	246	218
Manganese	[mg/kg WW]	3.2	18.4	18.0	18.8	59.3	31.3	27.6	28.2	25.1
Copper	[mg/kg WW]	3.5	2.9	2.9	3.2	4.1	7.2	9.4	3.1	3.2
Zinc	[mg/kg WW]	7.7	19.7	16.8	17.0	45.7	39.3	34.0	12.6	10.8
Nickel	[mg/kg WW]	0.31	0.28	0.28	0.28	0.75	0.38	0.45	1.10	1.03
Molybdenum	[mg/kg WW]	0.21	0.23	0.21	0.21	0.27	0.39	0.37	0.41	0.38
Cobalt	[mg/kg WW]	0.15	0.14	0.13	0.14	0.16	0.14	0.13	0.30	0.27

Table 3 (continued).

Biogas plant		P5 MD	P5 SD	P5 RDS	P6 MD	P6 SD	P6 RDS	P7 MD	P7 SD	P7 RDS	P8 MD	P8 SD	P8 RDS
Sampling		Feb 3, 2014			July 23, 2014						July 29, 2014		
DW	[%WW]	8.5	6.1	5.0	5.9	6.2	5.7	7.4	6.7	4.9	6.6	5.5	2.7
oDW	[%DW]	81.8	76.6	75.9	72.4	72.9	73.3	78.0	75.5	71.9	70.6	71.2	73.2
Calcium	[mg/kg DW]	12, 500	16, 800	17, 100	46, 492	52, 162	46, 315	28, 067	27, 481	26, 537	37, 429	44, 625	39, 775
Potassium	[mg/kg DW]	39, 500	56, 700	63, 400	69, 652	73, 184	78, 760	59, 921	59, 297	62, 492	50, 220	56, 154	50, 564
Magnesium	[mg/kg DW]	5, 300	6, 300	6, 200	10, 692	11, 172	10, 294	11, 246	10, 895	9, 831	12, 019	13, 857	11, 057
Iron	[mg/kg DW]	1, 200	1, 200	1, 200	2, 106	2, 510	3, 279	2, 260	2515	2, 698	3, 268	2, 651	6, 337
Sodium	[mg/kg DW]	3, 500	4, 500	3, 800	4, 620	4, 168	5, 086	7, 964	9, 191	10, 519	4, 911	5, 473	4, 843
Manganese	[mg/kg DW]	178	233	234	251	249	246	299	339	374	314	365	422
Copper	[mg/kg DW]	24	36	32	122	111	110	55	62	76	285	286	377
Zinc	[mg/kg DW]	210	300	380	294	205	462	340	382	463	308	374	486
Nickel	[mg/kg DW]	2.26	3.04	2.69	3.93	4.43	3.75	6.82	7.14	8.71	9.14	8.67	-
Molybdenum	[mg/kg DW]	1.28	1.74	1.56	2.25	2.55	2.30	3.65	3.97	4.33	2.64	3.29	10.53
Cobalt	[mg/kg DW]	1.12	1.59	1.70	1.18	1.19	1.35	1.60	1.85	2.43	2.31	2.04	4.52
Selenium	[mg/kg DW]	0.89	1.00	1.70	0.62	0.58	0.72	1.32	1.60	1.82	0.88	1.20	1.32
Calcium	[mg/kg WW]	1, 063	1, 025	855	2, 743	3, 234	2, 640	2, 077	1, 841	1, 300	2, 470	2, 454	1, 074
Potassium	[mg/kg WW]	3, 358	3, 459	3, 170	4, 109	4, 537	4, 489	4, 434	3, 973	3, 062	3, 315	3, 088	1, 365
Magnesium	[mg/kg WW]	451	384	310	631	693	587	832	730	482	793	762	299
Iron	[mg/kg WW]	102	73	60	124	156	187	167	168	132	216	146	171
Sodium	[mg/kg WW]	298	275	190	273	258	290	589	616	515	324	301	131
Manganese	[mg/kg WW]	15.2	14.2	11.7	14.8	15.5	14.0	22.1	22.7	18.3	20.7	20.1	11.4
Copper	[mg/kg WW]	2.1	2.2	1.6	7.2	6.9	6.3	4.1	4.1	3.7	18.8	15.7	10.2
Zinc	[mg/kg WW]	17.9	18.3	19.0	17.3	12.7	26.3	25.1	25.6	22.7	20.4	20.6	13.1
Nickel	[mg/kg WW]	0.19	0.19	0.13	0.23	0.27	0.21	0.50	0.48	0.43	0.60	0.48	-
Molybdenum	[mg/kg WW]	0.11	0.11	0.08	0.13	0.16	0.13	0.27	0.27	0.21	0.17	0.18	0.28
Cobalt	[mg/kg WW]	0.10	0.10	0.09	0.07	0.07	0.08	0.12	0.12	0.12	0.15	0.11	0.12

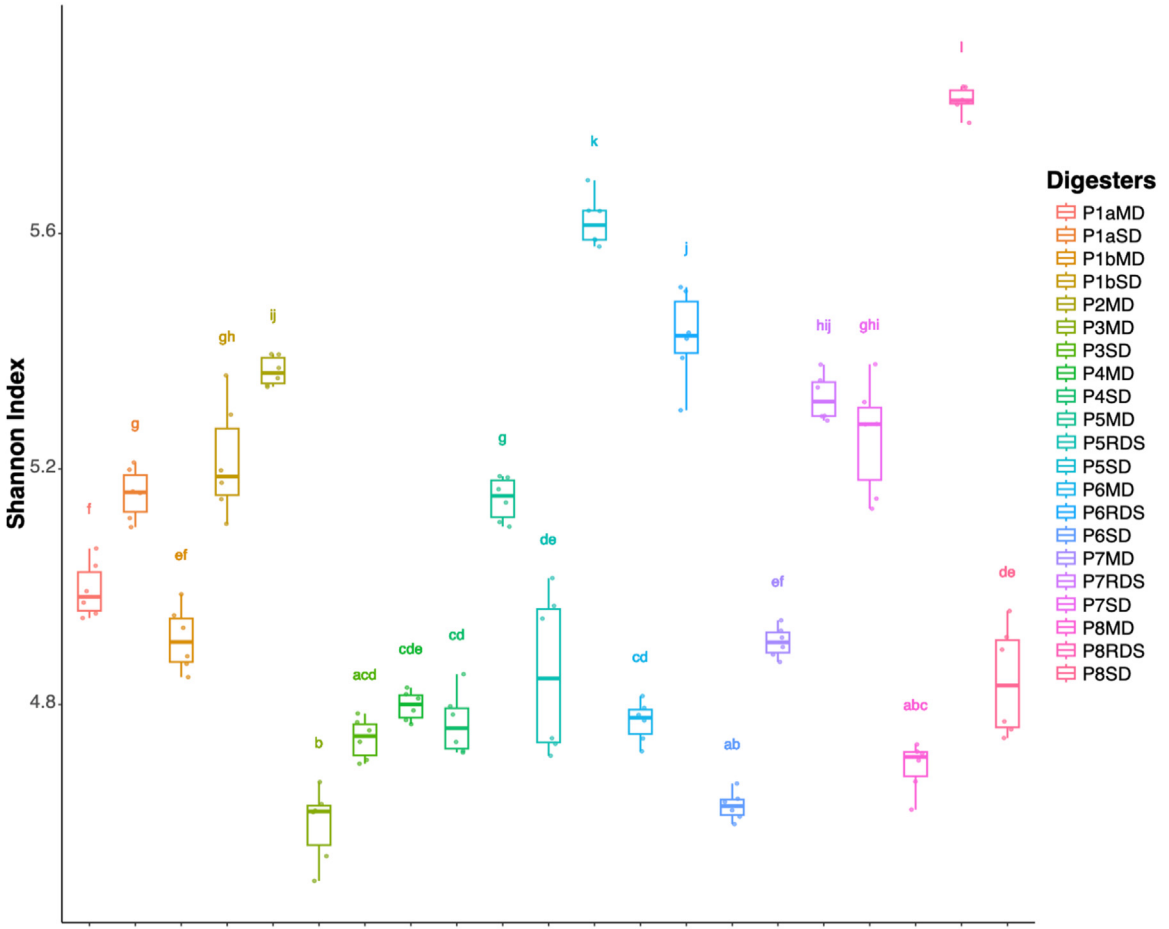

Fig. 1. Shannon diversity index values in different digesters of sampled BPs (P1-P8, see Table 1 for details).

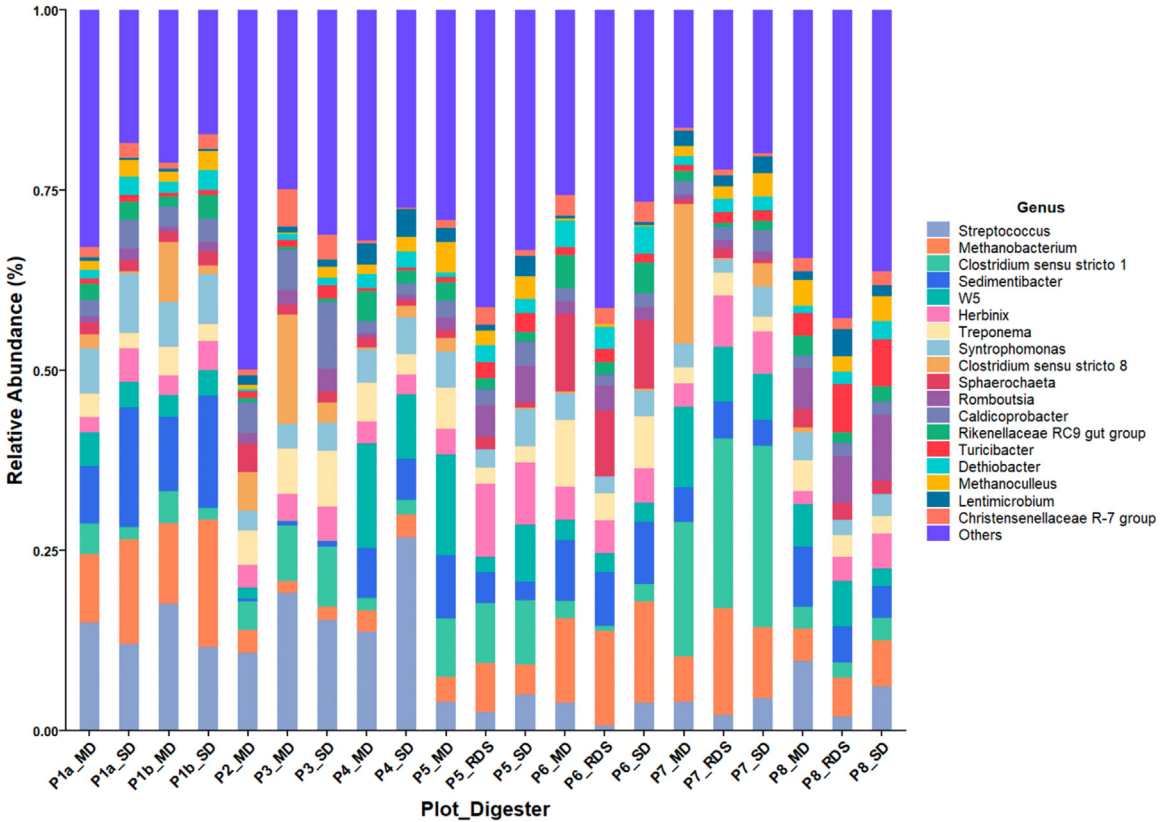

Fig. 2. Taxonomic profiles in the various digester types of BPs (see Table 1 for details) at genus level. The abundance values were normalized and expressed as a percentage of the total abundance of all identified species. The top 20 genera are displayed while the remaining genera were grouped together as "Others".

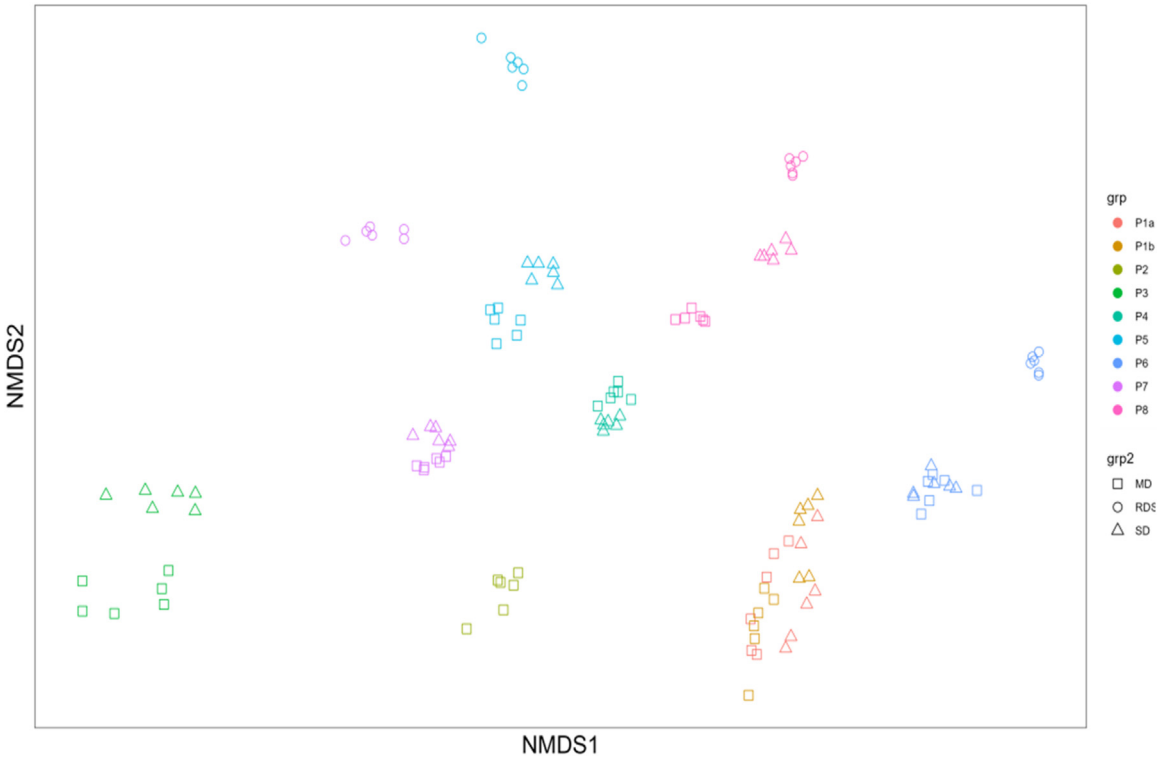

Fig. 3. Euclidean non-metric multidimensional scaling (NMDS; Stress: 0.103) with centered Log-ratio (CLR) transformation in the various digester types of BPs (see Table 1 for details) at the ASV level.

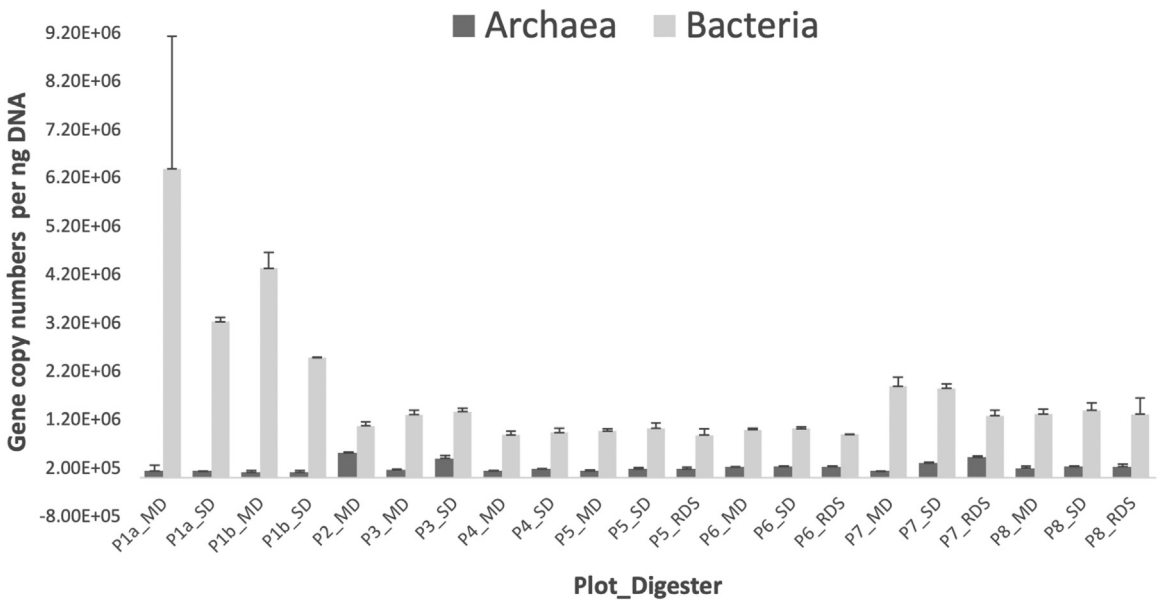

Fig. 4. qPCR data in different digesters of sampled BPs (P1-P8, see table 1 for details).

4. Experimental Design, Materials and Methods

4.1. Sampling and DNA-extraction

In total, eight biogas plants (BPs) were sampled at different times in June 2012 (P1), December 2012 (P2), May 2013 (P3), December 2013 (P4), February 2014 (P5), and July 2014 (P6, P7, P8). The biogas plant P1 consisted of two separate lines of fermenters, marked as P1a, P1b. In the biogas plants P1, P3, P4 and P5, main digesters (MD) and secondary digesters (SD) were sampled. In the biogas plant P2, only the main digester (MD) was sampled. In addition to main and secondary digesters, the residue digester storage (RDS) was also sampled at the biogas plants P5 to P8. Samples were pre-treated before DNA extraction as described before [3]. In brief, large particles in substrates and slurries from digesters were removed by washing with sterile phosphate buffered saline (PBS), pH 7.4, by centrifugation at 500 $\times$ *g* for 10 min. The supernatants were further processed at 11, 800 $\times$ *g* for 30 min, followed by washing with PBS and 0.85 % (wt./vol.) sterile KCl solution. For silage samples, microbial cells were extracted from 100-g portions suspended in 700 ml PBS at 4°C for 40 min, removing large plant material with a 2-mm sieve before continuing with 500 $\times$ *g* centrifugation for 10 min. Nucleic acids were finally using a phenol-chloroform extraction method described elsewhere (Method 5b, [4]).

4.2. Process parameter analyses

The digester material obtained from the biogas plant underwent a comprehensive analysis following the applicable DIN EN protocols. Specifically, the analysis covered the dry matter content, as per DIN EN 12880 [5] at 105°C in the drying oven, the organic dry matter content, in accordance with DIN EN 12879 [6] by incineration at 550°C in the muffle furnace, the pH level, adhering to DIN EN 15933 [7], and the NH_4^+-N content, as specified by DIN EN 38406–5 [8]. In addition, the ratio of volatile organic acids (VOA) to total inorganic carbon (TIC) is determined with Titroline 5000 from SI Analytics by using 20mL of a centrifuged sample and titrating with sulfuric acid (0.1N / 0.05mol/L) for the TIC value up to pH 5.0 and for the VOA value up to pH 4.4. The content of manganese, copper, nickel, molybdenum, and cobalt were measured by graphite furnace atomic absorption spectroscopy (AAS) (ZEEnit 650, Analytik Jena AG, Jena, Germany). Samples were diluted in 2.8 % (wt./vol.) HNO3. A solution of 1 % (wt./vol.) $NH_4H_2PO_4$ and 0.1 % (wt./vol.) Mg $(NO_3)_2$ was used as matrix-modifier. Carboxylic acids were measured by high-performance liquid chromatography (HPLC) and UV detection (UltiMate 3000, Dionex Thermo Fisher Scientific Inc., Sunnyvale, CA.). Filtered sample material was acidified with methylsulfonic acid to a final concentration of 27 mM and centrifuged.

4.3. Processing of raw 16S rRNA gene data

A total of 126 nucleic acid samples collected from the different BPs were amplified by PCR using the primer set Primers 515F and 909R (for details see [9]) and the resulting PCR products were sent to a sequencing facility (Starseq, Mainz, Germany) for 300-bp paired-end sequencing with V3 chemistry on an Illumina MiSeq system. The DADA2 pipeline version 1.16 [1] in R version 4.1.0 [10] was employed to process the raw 16S rRNA gene sequences. In brief, primers and barcodes were removed, followed by trimming forward and reverse reads at positions 260 and 150 from the end or any position with a trunQ-score of 2. Error models were constructed with random sample picking. The resulting two separate amplicon sequence variant (ASV) tables of different runs were merged into one final ASV table. Subsequently, the ASV table was filtered by removing chimeras and ASVs were taxonomically classified using the SILVA reference version 138 [11].

4.4. Analysis of processed 16S rRNA gene data

ASV that were identified as mitochondrial or chloroplast sequences, or not classified into prokaryotic phyla, were excluded from the dataset. A total of 12, 693 ASV were identified, with the most abundant phyla being Firmicutes (53.97 %), Proteobacteria (12.07 %), Actinobacteriota (10.19 %), and Bacteroidota (9.26 %). ASVs with a relative abundance of less than 0.05 % were removed from the 16S rRNA gene dataset. Bayesian-Multiplicative replacement of count zeros (cmultRepl) in the R package zCompositions package version 1.4.0.1 [12] was employed to impute zeros in compositional count data. A centered log-ratio (CLR) transformation implemented using the R package compositions version 2.0-5 [13] was applied to account for variations in sequencing depth among samples. Differences in community composition were visualized using non-metric multidimensional scaling (NMDS) based on Aitchison distance, which was implemented in the R package vegan version 2.6-4 [14]. Shannon diversity indices were calculated in R by using the R packages vegan [14].

4.5. Quantitative PCR

Bacterial and archaeal 16S rRNA gene copy numbers were determined by real-time PCR following protocols described earlier [2,15]. In brief, the StepOnePlus Real-Time PCR system (Life Technologies GmbH) was used to perform 20 µl PCR reactions containing 2 µl template DNA, 500 nM of each primer (custom synthesized by Biomers, Ulm, Germany) and 200 nM of the respective probe (Biomers) in the Maxima Probe qPCR Master Mix (Thermo Fisher Scientific Inc., Waltham, MA) including ROX (carboxy-X-rhodamine). To detect Bacteria and Archaea probes were labelled with 5′-FAM (6-carboxyfluorescein) and 3′-BHQ1 (Blackhole Quencher 1), or 5′-6-TAMRA (carboxytetramethylrhodamine) and 3′-BHQ2 (Blackhole Quencer 2), respectively. The following PCR conditions were applied: 10 min at 95°C followed by 45 cycles at 95°C for 15 s and 60°C for 60 s. Signals were recorded at 60°C. The PCR efficiency was 95 % ($R^2 = 0.999$) and 101 % ($R^2 = 0.999$) for *Bacteria* and *Archaea*.

Limitations

The study is based on the examination of eight different agricultural biogas production plants in Federal State of Lower Saxony (Germany). These plants may not be fully representative for the whole state limiting extrapolation of the results beyond the plants investigated. For each plant only one sampling event was implemented, thus, temporal variation at each plant could not be evaluated. The composition of the microbiome was analysed with studying the 16S rRNA gene, which occurs in all bacteria and archaea. This gene can give indications of some functional potentials, i.e., methanogenesis. However, it cannot conclusively demonstrate general or specific metabolic activities, e.g. utilization of carbon or nitrogen sources or the production of secondary metabolites.

Ethics Statement

The authors have read and follow the ethical requirements for publication in Data in Brief and confirming that the current work does not involve human subjects, animal experiments, or any data collected from social media platforms.

Credit Author Statement

Sascha M.B. Krause: Validation, Writing-Reviewing and Editing, Formal analysis, Supervision, **Rui Wang:** Writing - Original Draft, Visualization, Formal analysis, **Anja B. Dohrmann:** Writing - Original Draft, Visualization, Formal analysis, **Achim Loewen:** Resources, Formal analysis, Investigation, **Meike Walz:** Resources, Formal analysis, Investigation, **Christoph C. Tebbe*:** Writing-Reviewing and Editing, Conceptualization, Funding acquisition.

Data Availability

Dataset: Abundance and composition of microbiomes in agricultural biogas plants in Lower Saxony, Germany, with variation in organic substrates, process parameters and nutrients. (Original data) (OpenAgrar).
F2_DADA2_ASV_data.xlsx (Original data) (OpenAgrar).

Acknowledgements

We thank all owners of the biogas plant analysed for providing the sample materials. We thank Karin Trescher, Jana Usarek and Britta Müller (Thünen Institute) for their excellent technical assistance. We would also like to acknowledge the support of Martin Bolz from the Thünen IT facility for the maintenance of the Linux-based computer systems. The work was supported by the Lower Saxony Ministry of Food, Agriculture, Consumer Protection and Regional Development, Hannover, Germany.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2024.111095.

References

[1] B.J. Callahan, P.J. McMurdie, M.J. Rosen, A.W. Han, A.J. Johnson, S.P. Holme, s, DADA2: High-resolution sample inference from Illumina amplicon data, Nat. Methods 13 (2016) 581–583, doi:10.1038/nmeth.3869.
[2] Y. Yu, C. Lee, J. Kim, S. Hwang, Group-specific primer and probe sets to detect methanogenic communities using quantitative real- time polymerase chain reaction, Biotechnol. Bioeng. 89 (2005) 670–679, doi:10.1002/bit.20347.
[3] A.B. Dohrmann, M. Walz, A. Loewen, C.C. Tebbe, Clostridium cluster I and their pathogenic members in a full-scale operating biogas plant, Appl. Microbiol. Biotechnol. 99 (2015) 3585–3598, doi:10.1007/s00253-014-6261-y.
[4] A.B. Dohrmann, S. Baumert, L. Klingebiel, P. Weiland, C.C. Tebbe, Bacterial community structure in experimental methanogenic bioreactors and search for pathogenic clostridia as community members, Appl. Microbiol. Biotechnol. 89 (2011) 1991–2004, doi:10.1007/s00253- 010-2955-y.
[5] DIN EN 12880, Characterization of Sludges—Determination of Dry Residue and Water Content, Beuth Verlag, Berlin, 2001.
[6] DIN EN 12879, Characterization of Sludges—Determination of Loss on Ignition of Dry Mass, Beuth Verlag, Berlin, 2001.
[7] DIN EN 15933, Sludge, treated biowaste and soil—determination of pH. Beuth Verlag, Berlin, 2012.

[8] DIN EN 38406–5, German Standard Methods for the Examination of Water, Waste Water and Sludge; Cations (group E); Determination of Ammonia-Nitrogen (E5), Beuth Verlag, Berlin, 1983.
[9] J. Li, J. Rui, Z. Pei, X. Sun, S. Zhang, Z. Yan, Y. Wang, X. Liu, T. Zheng, X. Li, Straw- and slurry-associated prokaryotic communities differ during co-fermentation of straw and swine manure, Appl. Microbiol. Biotechnol. 98 (2014) 4771–4780, doi:10.1007/s00253-014-5629-3.
[10] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2021 URL https://www.R-project.org/.
[11] M.R McLaren, B.J. Callahan, Silva 138.1 prokaryotic SSU taxonomic training data formatted for DADA2 [Data set] (2021) Zenodo. doi:10.5281/zenodo.4587955.
[12] J. Palarea-Albaladejo, J. Martín-Fernández, ZCompositions - R package for multivariate imputation of left-censored data under a compositional approach, Chemometr. Intell. Lab. 143 (2015) 85–96, doi:10.1016/j.chemolab.2015.02.019.
[13] K.G. van den Boogaart, R. Tolosana-Delgado, M. Bren, compositions: compositional data analysis. R Package Version 2.0-5, 2022.
[14] J. Oksanen, F.G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P.R. Minchin, R.B. O'Hara, G.L. Simpson, P. Solymos, M.H.H. Stevens, E. Szoecs, H. Wagner, vegan: community ecology package. R Package Version 2.6-4, 2017.
[15] M. Hemkemeyer, B.T. Christensen, R. Martens, C.C. Tebbe, Soil particle size fractions harbour distinct microbial communities and differ in potential for microbial mineralisation of organic pollutants, Soil Biol. Biochem. 90 (2015) 255–265, doi:10.1016/j.soilbio.2015.08.018.