Supplementary Information

Title: pH and organic carbon dose rates control microbially-driven bioremediation efficacy in alkaline bauxite residue

Authors: Talitha C. Santini, Laura I. Malcolm, Gene W. Tyson, Lesley A. Warren

Contents: Methods and Materials, 1 figure, 12 tables, 29 pages

Materials and Methods

Statistical analyses were performed using GenStat (geochemical data; Genstat Release 17.1.0.14713; VSN International). Repeated measures ANOVA was performed on geochemical data (pH, VFA concentrations, CO₂ and CH₄ production) collected during the bioreactor experiments, with initial pH, initial EC, and organic carbon dose rates and sterility (live or sterile) as factors. Tukey's Honestly Significant Difference (HSD) was used as a post hoc test to separate means unless otherwise stated. Interactions between factors were not significant unless otherwise stated. A significance level of $\alpha = 0.05$ was used throughout. Repeated measures ANCOVA was used to check for the effect of initial pH (time = 0 days) as a covariate potentially contributing to the significance of salinity and organic carbon as factors controlling pH neutralization.

Rates and extents of pH neutralization in inoculated treatments were compared using several methods beyond the comparison of mean treatment pH values in repeated measures ANOVA. Minimum pH, and time to minimum pH, in each treatment was calculated as the time (days) at which the minimum pH was observed in each replicate. After transforming pH data to H^+ concentration in suspensions, the extent of pH neutralization in each replicate was calculated as the difference in H^+ concentration between time = 0 days and the time at which minimum pH was observed, and the maximum rate of pH neutralization observed in each replicate over each 24 hour time period. Minimum pH, time to minimum pH, extent of pH neutralization, and maximum rate of pH neutralisation were compared between treatments by ANOVA, with two sample *t*-tests or Tukey's HSD used as post hoc tests to separate means as required. Time to equilibrium pH in each treatment was calculated as the earliest day on which there was no significant difference between the observed and final (day 16) pH,

S2

according to the least significant difference of means (for interaction of time x initial pH x initial salinity x organic carbon dose rate) returned by repeated measures ANOVA. All gaseous concentration data were corrected for the effect of daily aqueous sample removals. Carbon recovery index, representing an overall conversion efficiency of glucose plus organic carbon in yeast extract and peptone to fermentation products, was calculated from the sum of total $CO_{2(g)}$ (calculated by integrating $CO_{2(g)}$ production rates over time between collection of samples), acetic acid, and ethanol produced. A combined yield coefficient for major acidic products (mmol of acetic acid and $CO_{2(g)}$ produced per mmol of organic carbon supplied in glucose, yeast extract and peptone) was also calculated.¹ Calculations for both carbon recovery index and yield coefficient of acidic products used manufacturer's specifications for carbohydrate content, assuming all carbohydrates were present as glucose.

Microbial DNA extraction, sequencing, and statistical analyses

Microbial DNA was extracted from frozen samples (MoBio PowerSoil DNA Isolation kit, MoBio, Carlsbad, CA) according to manufacturer's protocols, and PCR amplification was performed using the Q5 Hot Start High-Fidelity 2X Master Mix Kit (New England Biolabs, Ipswich, MA) and modified versions of the universal (Archaea and Bacteria) primers 926F (5'-AAACTYAAAKGAATTGRCGG-3') and 1392R (5'-ACGGGCGGTGWGTRC-3'), targeting the V6-V8 region.² Sequencing of archaeal and bacterial 16S rRNA genes was performed using the Illumina MiSeq platform and reagents according to manufacturer's protocols. Microbial DNA sequencing data was processed using QIIME (v 1.8.0).³ Multiplex identifiers and primers were removed, sequences containing ambiguous base calls, less than 150 bp in length, or containing homopolymer runs >6 bp were discarded, as were chimeric sequences, and operational taxonomic units (OTUs) were defined by clustering at 97%

S3

similarity using an open reference OTU picking strategy. After sequence alignment, phylogenetic trees were created with FastTree⁴ and taxonomy was assigned to OTUs using BLAST against a curated GreenGenes database.⁵

After trimming and quality filtering, an average of 143706 reads were returned per sample; these were rarefied to a uniform depth of 20000 reads per sample. Relative abundances of OTUs were corrected for differences in 16S rRNA gene copy number using CopyRighter (v 0.46)⁶ before analysis of alpha and beta diversity. Alpha diversity was compared using Shannon (H'), reciprocal Simpson, chao1, and Faith's Phylogenetic Diversity metrics. Community composition in each sample was visualized by nonmetric multidimensional scaling based on Bray-Curtis distance matrices. PERMANOVA⁷ and PERMDISP,^{8,9} implemented in PRIMER (v 7.0.10, with PERMANOVA+ v 1 add-in; PRIMER-E, Plymouth UK),¹⁰ were used to test for statistically significant differences in community composition and dispersion between treatments based on Bray-Curtis distance matrices. Permutation *P*-values were used unless low unique permutations necessitated the use of Monte Carlo asymptotic *P*-values. Significant environmental drivers of community composition were identified using distance-based multivariate multiple regression (DistLM), implemented in PRIMER, with a forward selection procedure using 9999 permutations.

References for Supplementary Information Materials and Methods

(1) El-Mansi, E. M. T.; Bryce, C. F. A.; Demain, A. L.; Allman, A. R.; Sanchez, S.; Dahhou, B., Eds. *Fermentation Microbiology and Biotechnology*, 3rd ed.; CRC Press: Boca Raton FL, 2013.

(2) Matsuki, T.; Watanabe, K.; Fujimoto, J.; Miyamoto, Y.; Takada, T.; Matsumoto, K.; Oyaizu, H.; Tanaka, R. Development of 16S rRNA gene-targeted group specific primers for the detection and identification of predominant bacteria in human feces. *Appl. Environ. Microbiol.* **2002**, *68*, 5445-5451.

(3) Caporaso, J. G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F. D.; Costello,
E. K.; Fierer, N.; Gonzalez Peña, A.; Goodrich, J. K.; Gordon, J. I.; Huttley, G. A.; Kelley, S.
T.; Knights, D.; Koenig, J. E.; Ley, R. E.; Lozupone, C. A.; McDonald, D.; Muegge, B. D.;
Pirrung, M.; Reeder, J.; Sevinsky, J. R.; Turnbaugh, P. J.; Walters, W. A.; Widmann, J.;
Yatsunenko, T.; Zaneveld, J.; Knight, R. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 2010, *7*, 335-336.

(4) Price, M. N.; Dehal, P. S.; Arkin, A. P. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **2009**, *26*, 1641-1650.

(5) DeSantis, T. Z.; Hugenholtz, P.; Larsen, N.; Rojas, M.; Brodie, E. L.; Keller, K.; Huber, T.; Hu, P.; Andersen, G. L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **2006**, *72*, 5069-5072.

(6) Angly, F. E.; Dennis, P. G.; Skarshewski, A.; Vanwonterghem, I.; Hugenholtz, P.; Tyson, G. W. CopyRighter: a rapid tool for improving the accuracy of microbial community profiles through lineage-specific gene copy number correction. *Microbiome* **2014**, *2*, 11.

(7) Anderson, M. J. A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* **2001**, *26*, 32-46.

(8) Anderson, M. J. Distance-based tests for homogeneity of multivariate dispersions. *Biometrics* **2006**, *62*, 245-253. (9) Anderson, M. J.; Ellingsen, K. E.; McArdle, B. H. Multivariate dispersion as a measure of beta diversity. *Ecol. Lett.* **2006**, *9*, 683-693.

(10) Clarke, K. R.; Gorley, R. N. *PRIMER v7: User Manual/Tutorial*; PRIMER-E: Plymouth, UK, 2015.

Supplementary Information Figure 1. Bauxite residue pH in sterile treatments, at (a) high initial pH; and (b) low initial pH. Bauxite residue and soil were sterilized by gamma irradiation at 50 kGy. Values displayed are the mean of three replicates; error bars indicate \pm 1 standard error of the mean.



Supplementary Information Table 1. Geochemical data for bioreactors during operation. Ethanol:CO₂ and acetic acid:CO₂ ratios determined

in final samples (day 16), based on molar ratios.

			Time to minimum	Extent of pH	Fastest rate of pH change	Total CO2	Yield coefficient	Carbon recovery		
-	-	Minimum	pH	neutralisation	(µmol H ⁺	yield	of acidic	index	Ethanol:	Acetic:
Treatment	Replicate	рН	(days)	(µmol H ⁺)	day")	(mmol)	products	(%)	CO ₂	CO2
High salinity, high pH, high C dose	1	6.95	10	0.11	0.07	15.65	0.30	24.71	2.85	0.63
High salinity, high pH, high C dose	2	7.01	16	0.10	0.06	18.70	0.29	20.54	1.96	0.31
High salinity, high pH, high C dose	3	6.95	10	0.11	0.08	11.60	0.26	20.05	3.00	0.86
High salinity, high pH, low C dose	1	9.65	16	0.00022	0.00	0.48	0.059	1.79	-	4.27
High salinity, high pH, low C dose	2	9.68	16	0.00021	0.00	0.22	0.083	2.69	-	14.98
High salinity, high pH, low C dose	3	7.32	16	0.05	0.02	15.87	0.82	34.52	1.08	1.19
High salinity, low pH, high C dose	1	5.13	7	7.41	4.38	25.86	0.72	30.27	1.11	1.35
High salinity, low pH, high C dose	2	5.44	10	3.63	1.16	37.51	0.79	29.90	0.73	0.78
High salinity, low pH, high C dose	3	5.97	10	1.07	0.56	37.61	0.84	39.19	1.26	0.87
High salinity, low pH, low C dose	1	6.60	3	0.25	0.24	16.64	1.19	49.19	1.24	2.03
High salinity, low pH, low C dose	2	6.30	10	0.50	0.35	14.40	0.82	26.95	0.48	1.40
High salinity, low pH, low C dose	3	6.85	16	0.14	0.11	21.18	0.85	21.63	0.09	0.71
Low salinity, high pH, high C dose	1	6.54	16	0.29	0.17	53.56	0.70	28.93	0.76	0.10
Low salinity, high pH, high C dose	2	6.16	16	0.69	0.38	47.78	0.93	22.16	0.03	0.64
Low salinity, high pH, high C dose	3	6.59	16	0.26	0.17	41.26	0.56	25.93	0.94	0.14
Low salinity, high pH, low C dose	1	7.02	10	0.10	0.07	4.24	0.32	20.48	3.42	2.24
Low salinity, high pH, low C dose	2	6.37	3	0.43	0.43	4.72	0.39	35.29	6.53	2.49
Low salinity, high pH, low C dose	3	6.73	10	0.19	0.12	16.51	0.58	27.97	1.16	0.49
Low salinity, low pH, high C dose	1	5.30	9	5.01	2.02	24.59	0.61	27.27	1.20	1.10
Low salinity, low pH, high C dose	2	6.02	3	0.95	0.65	53.57	0.98	39.44	0.82	0.53

Low salinity, low pH, high C dose	3	5.31	9	4.90	1.51	45.43	0.81	30.24	0.69	0.49
Low salinity, low pH, low C dose	1	5.39	11	4.07	3.98	9.26	0.81	33.14	1.34	2.72
Low salinity, low pH, low C dose	2	6.32	3	0.48	0.45	25.73	1.20	37.20	0.36	0.98
Low salinity, low pH, low C dose	3	6.56	3	0.27	0.25	16.79	0.91	28.81	0.38	1.30

- : indicates no ethanol produced

Supplementary Information Table 2. Repeated measures ANOVA results for bauxite residue bioreactors during incubation. Degrees of freedom (d.f.) correction factor for repeated measures is 0.1676. SS indicates sum of squares, F indicates F value. Bonferroni post-hoc test used a comparison-wise error rate of 0.0018 to correct for multiple comparisons. Significant differences between mean treatment pH values are indicated with lower-case letters.

Source of variation	d.f.	SS	F	<i>P</i> -value
Time	16	299.21	112.04	< 0.001
Time x salinity	16	17.59	6.59	0.001
Time x pH	16	35.59	13.33	< 0.001
Time x C dose	16	7.67	2.87	0.05
Time x salinity x pH	16	12.28	4.60	0.009
Time x salinity x C dose	16	12.48	4.67	0.008
Time x pH x C dose	16	4.31	1.61	0.20
Time x salinity x pH x C dose	16	11.28	4.23	0.01
Residuals	256	42.73		
Totals	407	1064.07		

Bonferroni post-hoc test for separation of means

Treatment	Mean treatment pH
Low pH, low salinity, high organic carbon	6.329 a
Low pH, high salinity, high organic carbon	6.389 ab
Low pH, low salinity, low organic carbon	7.380 abc
Low pH, high salinity, low organic carbon	7.516 abc
High pH, low salinity, high organic carbon	7.590 bc
High pH, low salinity, low organic carbon	8.000 c
High pH, high salinity, high organic carbon	8.551 c

Supplementary Information Table 3. Repeated measures ANCOVA results for high initial pH treatments in bauxite residue bioreactors during incubation. Degrees of freedom (d.f.) correction factor for repeated measures is 0.1116. SS indicates sum of squares, F indicates F value.

Source of variation	d.f.	SS	F	<i>P</i> -value
Time	16	210.69	79.07	< 0.001
Time x salinity	16	23.28	8.74	0.004
Time x C dose	16	6.20	2.33	0.137
Time x salinity x C dose	16	22.75	1.42	0.004
Residuals	128	21.31		
Totals	203	522.81		

Supplementary Information Table 4. Total fermentation product yields in bioreactors after 16 days of operation.

		Concentration (mmol L ⁻¹)									
								Iso-		Iso-	
Treatment	Replicate	Ethanol	Propanol	Butanol	Acetic acid	Propionic acid	Butyric acid	butyric acid	Valeric acid	valeric acid	Hexanoic acid
High salinity, high pH, high C dose	1	44.6	-	-	9.93	3.38	0.05	0.12	-	0.01	-
High salinity, high pH, high C dose	2	36.6	-	-	5.87	16.6	0.01	0.05	-	-	-
High salinity, high pH, high C dose	3	34.8	-	-	10.0	16.2	-	0.06	-	-	-
High salinity, high pH, low C dose	1	-	-	-	2.04	15.6	-	0.08	-	0.06	-
High salinity, high pH, low C dose	2	-	-	-	3.32	0.36	-	0.17	-	0.18	-
High salinity, high pH, low C dose	3	17.2	0.03	-	18.9	0.22	-	0.16	-	0.12	-
High salinity, low pH, high C dose	1	28.6	-	-	34.9	1.97	0.02	0.06	-	0.00	-
High salinity, low pH, high C dose	2	27.5	-	-	29.2	14.3	0.10	0.13	-	0.04	-
High salinity, low pH, high C dose	3	47.3	1.00	0.02	32.8	32.6	0.37	0.14	0.01	0.05	-
High salinity, low pH, low C dose	1	20.6	-	-	33.7	18.0	0.41	0.22	0.02	0.16	-
High salinity, low pH, low C dose	2	6.93	1.15	-	20.2	26.5	0.19	0.21	-	0.10	-
High salinity, low pH, low C dose	3	1.92	-	-	15.0	24.3	0.25	0.15	0.10	0.13	-
Low salinity, high pH, high C dose	1	40.7	7.75	0.13	5.53	18.7	0.58	0.18	0.03	0.16	-

		Concentration (mmol L ⁻¹)									
Treatment	Poplicata	Fthanol	Pronanal	Butanol	Acetic	Propionic	Butyric	Iso- butyric	Valeric	Iso- valeric	Hexanoic
Low salinity, high pH, high C dose	2	1.63	1.63	-	30.4	56.2	0.28	0.10	0.07	0.10	-
Low salinity, high pH, high C dose	3	38.9	0.46	1.17	5.91	0.35	2.31	0.09	0.13	0.09	0.13
Low salinity, high pH, low C dose	1	14.5	0.17	-	9.49	13.5	0.27	0.18	0.01	0.15	-
Low salinity, high pH, low C dose	2	30.8	0.27	0.08	11.8	7.36	0.67	0.28	0.02	0.21	-
Low salinity, high pH, low C dose	3	19.2	0.41	-	8.16	14.6	0.17	0.19	-	0.14	-
Low salinity, low pH, high C dose	1	29.5	-	-	27.0	6.69	0.03	0.02	-	-	-
Low salinity, low pH, high C dose	2	44.1	0.26	-	28.6	12.4	0.25	0.10	0.03	0.04	-
Low salinity, low pH, high C dose	3	31.3	-	-	22.3	3.75	0.02	0.06	-	0.02	-
Low salinity, low pH, low C dose	1	12.4	0.18	-	25.2	10.3	0.42	0.19	0.14	0.11	-
Low salinity, low pH, low C dose	2	9.29	0.58	-	25.2	13.7	0.25	0.34	0.01	0.08	-
Low salinity, low pH, low C dose	3	6.43	0.31	-	21.9	11.5	0.33	0.34	0.06	0.11	-

- : indicates analyte below detection limit.

Supplementary Information Table 5. Alpha diversity metrics for bauxite residue

suspension samples during bioreactor operation, and initial soil inoculants.

				Reciprocal		Faith's
Day	Treatment	Replicate	Shannon	Simpson	Chao1	PD
	Soil inoculant	1	11.74	522.62	9751	0.55
	Soil inoculant	2	11.71	517.30	9886	0.68
	Soil inoculant	3	11.78	516.28	10034	0.55
0	High salinity, high pH, high C dose	1	10.53	512.96	4143	0.17
	High salinity, high pH, high C dose	2	10.23	369.13	3688	0.33
	High salinity, high pH, high C dose	3	10.78	472.56	5361	0.17
	High salinity, high pH, low C dose	1	10.03	225.20	3338	0.51
	High salinity, high pH, low C dose	2	10.31	345.33	3806	0.53
	High salinity, high pH, low C dose	3	10.02	253.02	3774	0.89
	High salinity, low pH, high C dose	1	11.33	568.13	7650	0.79
	High salinity, low pH, high C dose	2	11.24	551.34	7551	0.68
	High salinity, low pH, high C dose	3	11.16	584.37	7117	0.62
	High salinity, low pH, low C dose	1	11.31	581.88	7395	0.62
	High salinity, low pH, low C dose	2	11.24	520.16	7449	0.69
	High salinity, low pH, low C dose	3	11.12	416.62	6913	0.68
	Low salinity, high pH, high C dose	1	11.61	337.03	9062	0.55
	Low salinity, high pH, high C dose	2	11.31	604.72	7242	0.55
	Low salinity, high pH, high C dose	3	10.90	654.10	5519	0.53
	Low salinity, high pH, low C dose	1	11.27	699.95	7067	0.53
	Low salinity, high pH, low C dose	2	11.23	613.01	6939	0.53
	Low salinity, high pH, low C dose	3	11.27	474.90	7613	0.71
	Low salinity, low pH, high C dose	1	10.97	469.44	7181	0.83
	Low salinity, low pH, high C dose	2	10.97	455.03	7009	0.70
	Low salinity, low pH, high C dose	3	11.07	525.61	7275	0.55
	Low salinity, low pH, low C dose	1	11.01	460.24	6984	0.71
	Low salinity, low pH, low C dose	2	10.96	476.70	6963	0.73
	Low salinity, low pH, low C dose	3	11.02	557.15	6895	0.70
3	High salinity, high pH, high C dose	1	2.50	1.77	1104	0.09
	High salinity, high pH, high C dose	2	1.76	1.41	960	0.07
	High salinity, high pH, high C dose	3	2.25	1.56	1204	0.07
	High salinity, high pH, low C dose	1	9.80	232.68	3153	0.85
	High salinity, high pH, low C dose	2	10.59	441.56	4814	0.84
	High salinity, high pH, low C dose	3	10.57	413.97	4585	0.67
	High salinity, low pH, high C dose	1	6.03	18.67	1473	0.65
	High salinity, low pH, high C dose	2	5.12	6.41	1295	0.55
	High salinity, low pH, high C dose	3	4.53	6.64	943	0.33
	High salinity, low pH, low C dose	1	3.54	3.46	1028	0.42
	High salinity, low pH, low C dose	2	4.86	7.15	1238	0.69
	High salinity, low pH, low C dose	3	5.52	15.58	982	0.29
	Low salinity, high pH, high C dose	1	4.47	4.61	1162	0.07

	Low salinity, high pH, high C dose	2	3.78	3.21	983	0.07
	Low salinity, high pH, high C dose	3	4.68	6.35	971	0.48
	Low salinity, high pH, low C dose	1	4.44	6.99	866	0.13
	Low salinity, high pH, low C dose	2	8.89	80.80	3185	1.75
	Low salinity, high pH, low C dose	3	4.84	4.63	1007	0.23
	Low salinity, low pH, high C dose	1	4.94	8.98	1162	0.82
	Low salinity, low pH, high C dose	2	5.72	8.83	1496	0.80
	Low salinity, low pH, high C dose	3	6.13	24.27	1459	1.45
	Low salinity, low pH, low C dose	1	6.39	24.24	1582	1.13
	Low salinity, low pH, low C dose	2	5.90	13.81	1468	1.03
	Low salinity, low pH, low C dose	3	6.11	11.50	2110	1.04
16	High salinity, high pH, high C dose	1	3.68	2.61	753	1.03
	High salinity, high pH, high C dose	2	6.38	14.50	1054	1.50
	High salinity, high pH, high C dose	3	7.84	58.59	1764	2.10
	High salinity, high pH, low C dose	1	1.69	1.86	363	0.07
	High salinity, high pH, low C dose	2	4.20	7.21	576	0.95
	High salinity, high pH, low C dose	3	5.02	10.48	549	1.98
	High salinity, low pH, high C dose	1	5.91	13.16	1526	0.50
	High salinity, low pH, high C dose	2	5.55	12.76	1163	0.45
	High salinity, low pH, high C dose	3	5.38	9.59	1336	1.14
	High salinity, low pH, low C dose	1	4.94	8.55	1247	0.82
	High salinity, low pH, low C dose	2	4.88	6.69	1160	1.32
	High salinity, low pH, low C dose	3	4.81	5.84	1474	1.56
	Low salinity, high pH, high C dose	1	5.23	10.98	1120	1.15
	Low salinity, high pH, high C dose	2	4.25	6.97	1001	1.33
	Low salinity, high pH, high C dose	3	4.95	7.69	1309	0.88
	Low salinity, high pH, low C dose	1	5.61	10.89	1436	1.23
	Low salinity, high pH, low C dose	2	3.07	2.37	723	1.02
	Low salinity, high pH, low C dose	3	5.29	11.69	1148	0.99
	Low salinity, low pH, high C dose	1	5.40	9.86	1346	0.67
	Low salinity, low pH, high C dose	2	5.99	17.29	1280	1.04
	Low salinity, low pH, high C dose	3	5.96	19.88	1660	1.12
	Low salinity, low pH, low C dose	1	6.66	24.78	1700	1.76
	Low salinity, low pH, low C dose	2	5.48	8.32	1652	0.96
	Low salinity, low pH, low C dose	3	5.12	6.95	1745	1.53

Supplementary Information Table 6. Relative abundances of the four key OTUs involved in fermentation-fuelled bioremediation as identified by SIMPER analysis and relative abundances. These OTUs comprised four of the top five OTUs explaining variation in microbial community structure between days 0 and 3 as identified by SIMPER analysis. The fifth OTU was associated with the soil inoculant (Solirubrobacterales sp.) and decreased substantially in relative abundance between days 0 and 3. Relative abundance is expressed as a percentage of total sequence reads, after correction for variations in 16S rRNA gene copy number between taxa (Angly et al., 2014). Values displayed are the mean of three replicates ± 1 standard error of the mean.

Day	Treatment	<i>Bacillaceae</i> sp. A	<i>Bacillaceae</i> sp. B	Bacillus sp.	Enterobacteriaceae sp.
0	High salinity, high pH, high C dose	0.0536 ± 0.0119	0.2326 ± 0.0273	1.4278 ± 0.2414	0.0104 ± 0.0015
	High salinity, high pH, low C dose	0.0399 ± 0.0071	0.2029 ± 0.0517	1.2829 ± 0.2716	0.0142 ± 0.0024
	High salinity, low pH, high C dose	0.0035 ± 0.0003	0.0421 ± 0.0039	0.0625 ± 0.0068	0.0029 ± 0.0023
	High salinity, low pH, low C dose	0.0064 ± 0.0009	0.0408 ± 0.0022	0.0496 ± 0.0028	0.0012 ± 0.0007
	Low salinity, high pH, high C dose	0.0026 ± 0.0006	0.0071 ± 0.0028	0.0464 ± 0.0072	0.0042 ± 0.0042
	Low salinity, high pH, low C dose	0.0026 ± 0.0009	0.0134 ± 0.0015	0.0573 ± 0.0033	0.0050 ± 0.0014
	Low salinity, low pH, high C dose	0.0035 ± 0.0006	0.0163 ± 0.0037	0.0539 ± 0.0109	0.0048 ± 0.0030
	Low salinity, low pH, low C dose	0.0035 ± 0.0006	0.0143 ± 0.0036	0.0475 ± 0.0019	0.0012 ± 0.0012
3	High salinity, high pH, high C dose	0.1304 ± 0.0300	85.847 ± 1.0519	5.2149 ± 0.4749	0.0718 ± 0.0140
	High salinity, high pH, low C dose	0.0987 ± 0.0138	0.5482 ± 0.0413	1.7801 ± 0.1389	0.5256 ± 0.0884
	High salinity, low pH, high C dose	1.3133 ± 0.0331	43.932 ± 8.3468	7.6355 ± 1.4102	15.682 ± 13.415
	High salinity, low pH, low C dose	3.1117 ± 1.7042	20.321 ± 5.4643	10.136 ± 4.7147	0.0469 ± 0.0371
	Low salinity, high pH, high C dose	5.2489 ± 3.1226	67.326 ± 6.8865	5.0225 ± 2.6480	0.0058 ± 0.0029
	Low salinity, high pH, low C dose	32.803 ± 15.269	4.9504 ± 1.9496	19.425 ± 6.5998	0.1376 ± 0.1165
	Low salinity, low pH, high C dose	0.5134 ± 0.1344	12.189 ± 3.2904	2.2316 ± 0.6060	43.087 ± 10.626
	Low salinity, low pH, low C dose	0.8456 ± 0.5497	14.569 ± 2.9832	3.6997 ± 0.8893	30.296 ± 12.198
16	High salinity, high pH, high C dose	0.1278 ± 0.0916	21.947 ± 21.339	0.2359 ± 0.1711	2.3219 ± 1.2751

High salinity, high pH, low C dose	0.5300 ± 0.2846	46.582 ± 26.475	0.1747 ± 0.0393	0.3281 ± 0.1249
High salinity, low pH, high C dose	1.5476 ± 0.4350	12.851 ± 3.9838	4.6621 ± 1.2515	47.780 ± 6.7723
High salinity, low pH, low C dose	0.3688 ± 0.0986	1.4624 ± 0.3605	1.5904 ± 0.3349	18.830 ± 10.146
Low salinity, high pH, high C dose	2.3670 ± 0.8192	10.259 ± 1.1097	1.7288 ± 0.7065	0.0399 ± 0.0042
Low salinity, high pH, low C dose	0.4429 ± 0.1974	0.3751 ± 0.2181	0.9305 ± 0.5814	0.3670 ± 0.2694
Low salinity, low pH, high C dose	0.3931 ± 0.2115	5.8065 ± 1.6073	2.2918 ± 1.2261	41.764 ± 9.453
Low salinity, low pH, low C dose	0.3781 ± 0.2693	3.4985 ± 1.4710	1.8470 ± 0.9402	36.348 ± 14.174

_

Supplementary Information Table 7. PERMANOVA results from analysis of Bray-Curtis dissimilarity matrix, with salinity, pH, organic carbon dose rate, and day as factors. *P*-values are based on 9999 permutations. Monte-Carlo asymptotic *P*-values are indicated by (MC). F values are pseudo-F determined by permutation. R^2 value can be interpreted as the proportion of variation explained.

Source of variation	d.f.	SS	F	R ²	<i>P</i> -value
Main effects					
Salinity	1	0.77643	3.391	0.015207	0.0001
рН	1	1.4761	6.447	0.034644	0.0001
C dose	1	0.60566	2.6452	0.010464	0.0003
Day	2	6.7365	14.711	0.1308	0.0001
Salinity x pH	1	1.0594	4.6269	0.046136	0.0001
Salinity x C dose	1	0.40185	1.755	0.0096044	0.0154
Salinity x day	2	0.88543	1.9335	0.017812	0.0007
pH x C dose	1	0.4705	2.0549	0.013418	0.003
pH x day	2	1.463	3.1948	0.041879	0.0001
C dose x day	2	0.76403	1.6684	0.012754	0.0041
Salinity x pH x C dose	1	0.51624	2.2546	0.031919	0.0014
Salinity x pH x day	2	1.2852	2.8066	0.068941	0.0001
Salinity x C dose x day	2	0.58654	1.2808	0.010717	0.0948
pH x C dose x day	2	0.66486	1.4519	0.017244	0.025
Salinity x pH x C dose x day	2	0.67929	1.4834	0.036892	0.023
Residuals	48	10.99		0.22897	
Totals	71	29.362			

Supplementary Information Table 8. PERMANOVA results from analysis of Bray-Curtis dissimilarity matrix, with salinity, pH, and organic carbon dose rate as factors defining *a priori* groups, separated by day. *P*-values are based on 9999 permutations. Monte-Carlo asymptotic *P*-values are indicated by (MC). F values are pseudo-F determined by permutation. R^2 value can be interpreted as the proportion of variation explained.

Day	Source of variation	d.f.	SS	F	\mathbf{R}^2	<i>P</i> -value
0	Main effects					
	Salinity	1	0.36931	3.3131	0.021486	0.0001
	рН	1	0.46137	4.139	0.029159	0.0001
	C dose	1	0.11521	1.0336	0.00031186	0.483
	Salinity x pH	1	0.33874	3.0388	0.037878	0.0001
	Salinity x C dose	1	0.11655	1.0455	0.00084616	0.4564
	pH x C dose	1	0.11946	1.0717	0.0013312	0.3869
	Salinity x pH x C dose	1	0.11712	1.0507	0.0018822	0.4408
	Residuals	16	1.7835		0.11147	
	Totals	23	3.4213			
	Contrasts	d.f.		<i>t</i> -statistic		<i>P</i> -value
	High salinity, high pH vs high salinity, low pH	8		2.2248		0.0017
	Low salinity, high pH vs low salinity, low pH	8		1.413		0.002
	High pH, high salinity vs high pH, low salinity	8		2.1064		0.0026
	Low pH, high salinity vs low pH, low salinity	8		1.2557		0.0039
3	Main effects	d.f.	SS	F	\mathbf{R}^2	<i>P</i> -value
	Salinity	1	0.67741	2.9449	0.37282	0.0001

pН	1	1.2018	5.2248	0.080985	0.0001
C dose	1	0.6831	2.9697	0.37757	0.0002
Salinity x pH	1	1.2263	5.3313	0.16605	0.0001
Salinity x C dose	1	0.47839	2.0797	0.041394	0.0004
pH x C dose	1	0.51777	2.2509	0.047958	0.0005
Salinity x pH x C dose	1	0.6753	2.9358	0.14842	0.0002
Residuals	16	3.6804		0.23003	
Totals	23	9.1406			
Contrasts	d.f.		<i>t</i> -statistic		<i>P</i> -value (MC)
High salinity, high pH, high C dose vs low salinity, high pH, high C dose	4		4.1561		0.0029
High salinity, high pH, low C dose vs low salinity, high pH, low C dose	4		1.5742		0.0875
High salinity, low pH, high C dose vs low salinity, low pH, high C dose	4		1.4119		0.1444
High salinity, low pH, low C dose vs low salinity, low pH, low C dose	4		1.3712		0.1474
High salinity, high pH, high C dose vs high salinity, low pH, high C dose	4		3.5735		0.0044
High salinity, high pH, low C dose vs high salinity, low pH, low C dose	4		2.0117		0.0289
Low salinity, high pH, high C dose vs low salinity, low pH, high C	4		1.8682		0.0426

	dose					
	Low salinity, high pH, low C dose vs low salinity, low pH, low C dose	4		1.3226		0.175
	High salinity, high pH, high C dose vs high salinity, high pH, low C dose	4		3.5073		0.004
	High salinity, low pH, high C dose vs high salinity, low pH, low C dose	4		1.1326		0.3098
	Low salinity, high pH, high C dose vs low salinity, high pH, low C dose	4		1.5719		0.0877
	Low salinity, low pH, high C dose vs low salinity, low pH, low C dose	4		0.81969		0.6241
16	Main effects	d.f.	SS	F	\mathbf{R}^2	<i>P</i> -value
	Salinity	1	0.61514	1.7809	0.022478	0.0033
	рН	1	1.276	3.6941	0.077546	0.0001
	C dose	1	0.57137	1.6542	0.01883	0.0212
	Salinity x pH	1	0.77956	2.2569	0.072358	0.0002
	Salinity x C dose	1	0.39344	1.1391	0.0080063	0.2874
	pH x C dose	1	0.49813	1.4422	0.025454	0.0711
	Salinity x pH x C dose	1	0.40311	1.1671	0.019234	0.257
	Residuals	16	5.5263		0.34541	
	Totals	23	10.063			
	Contrasts	d.f.		<i>t</i> -statistic		<i>P</i> -value
	High salinity, high pH vs high salinity, low pH	8		1.8067		0.0022

Low salinity, high pH vs low salinity, low pH	8	1.6433	0.0022
High pH, high salinity vs high pH, low salinity	8	1.5087	0.0034
Low pH, high salinity vs low pH, low salinity	8	1.3121	0.0438

Supplementary Information Table 9. PERMANOVA results from analysis of Bray-Curtis dissimilarity matrix based on comparison of *post hoc* groups within days. Group A contained initial soil inoculant (three replicates), Group B contained high salinity, high pH treatments, Group C contained all other treatments. *P*-values are based on 9999 permutations. Monte-Carlo asymptotic *P*-values are indicated by (MC). F values are pseudo-F determined by permutation. R^2 value can be interpreted as the proportion of variation explained.

Day	Source of variation	d.f.	SS	F	\mathbf{R}^2	<i>P</i> -value
0	Main effects					
	Post-hoc group	2	1.1867	5.0329	0.071316	0.0001
	Residuals	24	2.8294		0.11789	
	Totals	26	4.016			
	Contrasts	d.f.	Average distance between groups	<i>t</i> -statistic		<i>P</i> -value
	A vs B	7	0.71268	2.1130		0.0052 (MC)
	A vs C	19	0.55061	1.6167		0.0011
	B vs C	22	0.63929	2.6957		0.0001
3	Main effects	d.f.	SS	F	\mathbf{R}^2	<i>P</i> -value
	Post-hoc group	2	2.7301	4.2016	0.15602	0.0001
	Residuals	24	7.7973		0.32489	
	Totals	26	10.527			
	Contrasts	d.f.	Average distance between	<i>t</i> -statistic		<i>P</i> -value

			groups			
	A vs B	7	0.85064	1.7404		0.0243
	A vs C	19	0.96995	2.0228		0.0005
	B vs C	22	0.9728	2.1613		0.0001
16	Main effects	d.f.	SS	F	\mathbb{R}^2	<i>P</i> -value
	Post-hoc group	2	2.4301	3.1828	0.12499	0.0001
	Residuals	24	9.1621		0.38176	
	Totals	26	11.592			
	Contrasts	d.f.	Average distance between groups	<i>t</i> -statistic		<i>P</i> -value
	A vs B	7	0.99773	1.9172		0.0106 (MC)
	A vs C	19	0.98864	1.8527		0.0005
	B vs C	22	0.98450	1.6919		0.0011

Supplementary Information Table 10. PERMANOVA results from pairwise comparisons of days within each treatment based on Bray-Curtis dissimilarity matrix. Monte-Carlo asymptotic P-values (MC) are based on 9999 permutations. *t* values are for pairwise comparisons between days.

Treatment	Comparisons (time in days)	t	<i>P</i> -value (MC)
High salinity, high pH, high C dose	0 vs 3	4.2063	0.002
	0 vs 16	1.7801	0.0496
	3 vs 16	1.7268	0.1005
High salinity, high pH, low C dose	0 vs 3	1.1618	0.2943
	0 vs 16	1.8937	0.0392
	3 vs 16	1.8587	0.0443
High salinity, low pH, high C dose	0 vs 3	2.7951	0.0084
	0 vs 16	2.3938	0.0194
	3 vs 16	1.2041	0.2630
High salinity, low pH, low C dose	0 vs 3	2.3366	0.0186
	0 vs 16	2.2520	0.0186
	3 vs 16	1.2063	0.2751
Low salinity, high pH, high C dose	0 vs 3	2.8341	0.0074
	0 vs 16	1.8596	0.0455
	3 vs 16	1.5504	0.0962
Low salinity, high pH, low C dose	0 vs 3	1.7637	0.0559
	0 vs 16	2.2702	0.0211
	3 vs 16	1.4429	0.1199
Low salinity, low pH, high C dose	0 vs 3	2.3463	0.0165
	0 vs 16	2.3833	0.0160

	3 vs 16	0.5920	0.7869
Low salinity, low pH, low C dose	0 vs 3	2.2185	0.0220
	0 vs 16	1.9853	0.0318
	3 vs 16	0.7885	0.6313

Supplementary Information Table 11. Results of distance-based multivariate multiple regression (DistLM) based on Bray-Curtis dissimilarities for microbial community structures and measured environmental characteristics during bioreactor operation, using 9999 permutations under a forward selection procedure. Percentages listed with dbRDA axes indicate the percentage of variation explained by each axis out of the fitted model and total variation.

Environmental characteristic	<i>P</i> -value	Cumulative percentage of variation	Multiple partial dbRD	correlations with A axes
		explained (%)	Axis 1	Axis 2
			(66.41 % fitted, 17.33 % total)	(15.10 % fitted, 3.94 % total)
рН	0.0001	15.934	0.646	-0.691
Acetic acid concentration	0.0001	19.753	-0.650	-0.285
Daily $CO_{2(g)}$ production	0.0001	22.176	-0.156	-0.145
Ethanol concentration	0.0001	24.247	-0.321	-0.562
Daily rate of change in $H^{+}_{(aq)}$ concentration	0.45	26.096	-0.180	-0.323

Supplementary Information Table 12. PERMDISP results from analysis of Bray-Curtis dissimilarity matrix, with time (days) as group factor. *P*-values are based on 9999 permutations. Pseudo-F (determined by permutation) was 169.29 with d.f.₁ =2 and d.f.₂ =69. *t* values are for pairwise comparisons between levels of factor (day). Means and S.E. (standard errors) calculated from within-group dispersion.

Comparisons	t	<i>P</i> -value
Time 0 vs time 3	12.221	0.0001
Time 0 vs time 16	18.347	0.0001
Time 3 vs time 16	2.1119	0.0439
Within-group dispersion	Mean	S.E.
Within-group dispersion Time 0	Mean 0.37268	S.E. 0.012617
Within-group dispersion Time 0 Time 3	Mean 0.37268 0.61389	S.E. 0.012617 0.013178