Supplemental Information

Microbial diversity in engineered haloalkaline environments shaped by shared geochemical drivers observed in natural analogues

Talitha C. Santini[1,2,3,#], Lesley A. Warren[4], Kathryn E. Kendra[4]

[1] School of Geography, Planning, and Environmental Management, Steele Building, The University of Queensland, St Lucia QLD 4072

[2] Centre for Mined Land Rehabilitation, Sir James Foots Building, The University of Queensland, St Lucia QLD 4072

[3] School of Earth and Environment, The University of Western Australia, 35 Stirling Highway, Crawley WA 6009, Australia

[4] School of Geography and Earth Sciences, McMaster University, General Sciences Building, 1280 Main Street West, Hamilton, Ontario, L8S4K1, Canada

corresponding author: t.santini@uq.edu.au

17 pages, 8 figures, 8 tables

(Note: Supplemental Information Tables S7-S9 are provided as separate .xls files for ease of use)

Table S1. Additional geochemical properties of bauxite residue samples.

Sample name	Water extractable Ca (% wt)	Inorganic C (% wt)	Organic C (% wt)	Total N (% wt)	Exchangeable Ca (meq/100 g)	Exchangeable K (meq/100 g)	Exchangeable Mg (meq/100 g)	Exchangeable Na (meq/100 g)
BR1	0.02	20.9	-	-	2.00	0.17	-	14.83
BR2	-	4.19	-	-	3.95	0.51	-	25.99
BR3	-	0.99	-	-	8.04	-	-	20.50
BR4	0.10	0.77	4.68	0.2	6.15	1.03	5.16	1.42
BR5	-	0.72	17.4	1	-	1.03	-	71.13
BR6	-	0.04	0.24	-	3.03	-	-	0.08
BR7	-	0.48	0.09	-	4.50	-	-	0.86
BR8	0.03	2.02	2.61	-	10.74	-	0.81	-

- : below detection limit.

Table S2. Bacterial and fungal community diversity and richness in samples at phylum and OTU level, and correlations between diversity and richness metrics.

Sample name	Bacteria				Fungi				
	H'_{phy}	H'_{OTU}	L_{OTU}	S_{rar}	L_{phy}	H'_{phy}	H'_{OTU}	L_{OTU}	S_{rar}
BR1	1.28	4.68	0.98	215		-	-	-	-
BR2	1.17	4.49	0.98	180	0.00	0.00	2.97	0.91	47
BR3	1.20	4.57	0.98	240	0.00	0.00	1.75	0.79	7
BR4	0.75	3.15	0.91	103	-	-	-	-	-
BR5	1.16	5.31	0.99	522	0.41	0.79	4.14	0.96	95
BR6	1.66	6.33	1.00	750	0.09	0.22	4.65	0.98	126
BR7	1.65	6.70	1.00	1032	0.58	1.02	3.25	0.92	50
BR8	1.41	6.28	1.00	757	0.58	1.05	4.10	0.97	91
Correlations between metrics	Bacteria				Fungi				
	H'_{phy}	H'_{OTU}	L_{OTU}	S_{rar}	L_{phy}	H'_{phy}	H'_{OTU}	L_{OTU}	S_{rar}
L_{phy}						1.00***			
H'_{OTU}	0.94***				0.41	0.47			
L_{OTU}	0.87**	0.85**			0.50	0.55	0.97**		
S_{rar}	0.83*	0.95***	0.64		0.28	0.34	0.98***	0.91*	

Table S3. Correlations between UniFrac principal component loadings, chemical and physical properties of bauxite residues, and measures of bacterial community composition. Only statistically significant correlations are displayed. The relative abundances of phyla and classes within the bacterial community of each sample were used to calculate correlations with principal components.

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Chemical and physical properties							
EC	-0.8134*						
Exchangeable sodium percentage					-0.7610*		
pH					-0.7570*		
Total alkalinity				-0.7933*			
Moisture content			-0.7798*				
Water extractable Na	-0.8420**						
Gravel		0.8635**					
Sand	0.7069*						
Silt	-0.8317*						
Measures of bacterial community composition and diversity							
Clostridia				-0.8367**			
Actinobacteria			0.9184**				
Bacteroidetes		0.8743**					
Epsilonproteobacteria	-0.7069*						
Firmicutes	-0.7558*						
Gammaproteobacteria				-0.8020*			
Gemmatimonadetes	0.7274*						
Planctomycetes	0.8142*						
Proteobacteria			-0.7764*				
H'_{OTU}	0.8426**						
S_{rar}	0.9027**						

Table S4. Correlations between chemical and physical properties of bauxite residues, and measures of bacterial community composition. Only statistically significant correlations are displayed. The relative abundances of phyla and classes within the bacterial community of each sample were used to calculate correlations with chemical and physical properties.

Variable	Clay	EC	Ex Na	Gravel	Inorganic C	Moisture content	Organic C	Total alkalinity	Sand	Silt	W Na	pH
Alphaproteobacteria	-0.7493*					0.7454*						-0.8666**
Betaproteobacteria	0.8614**	0.7171*							-0.8275*			
Clostridia								0.9372***				
Deltaproteobacteria	0.7889*	0.7954*			0.8067*				-0.7716*		0.7330*	
Epsilonproteobacteria	0.9019**	0.9056**			0.7504*				-0.8492**	0.7871*	0.9075**	
Gammaproteobacteria							-0.7673*					
Acidobacteria			0.7707*				0.9146**					
Actinobacteria						-0.8411**						
Bacteroidetes				0.7750*								
Chloroflexi										-0.7510*		
Firmicutes	0.7535*	0.7504*								0.7516*	0.7780*	
Gemmatimonadetes												
OP11	0.8821**	0.9370***			0.8184*				-0.8313*	0.7757*	0.9333***	
Planctomycetes												
Proteobacteria						0.8413**						-0.7413*
Verrucomicrobia						0.7109*	0.8553**					
WS3	0.8811**	0.7551*							-0.8133*	0.7454*	0.7785*	
H'_{phy}						-0.7122*						

Ex Na: exchangeable Na

W Na: water extractable Na

Table S5. Correlations between UniFrac principal component loadings, chemical and physical properties of bauxite residues, and measures of fungal community composition. Only statistically significant correlations are displayed. The relative abundances of phyla and classes within the fungal community of each sample were used to calculate correlations with principal components.

Variable	PC1	PC2	PC3	PC4	PC5
Chemical and physical properties					
Total alkalinity	0.9535**				
Gravel content (%)		-0.8365*			
Moisture content (%)				-0.9875***	
Exchangeable Na				-0.9063*	
Organic C (%)				-0.8392*	
Measures of fungal community composition and diversity					
H'_{OTU}	-0.9172*				
L_{OTU}	-0.9749***				
S_{rar}	-0.8350*				
Dothideomycetes	0.8448*				
Cryptomycota				-0.8561*	
Dacrymycetes				-0.8534	

Table S6. Correlations between chemical and physical properties of bauxite residues, and measures of fungal community composition. Only statistically significant correlations are displayed. The relative abundances of phyla and classes within the fungal community of each sample were used to calculate correlations with chemical and physical properties.

	Asco	Basi	Chyt	Crypt	Dacry	Dothi	Glom	Sord	Total alkalinity	Moisture content	Ex Na	Organic C	pH
Asco			-0.833*			0.8852*	-0.909*	-0.816*					
Basi													-0.963**
Chyt													
Crypt					0.999***					0.888*	0.911*	0.989***	
Dacry										0.886*	0.910*	0.989***	
Dothi								-0.820*					
Glom													
Sord													
H'_{phy}	-0.977***					-0.936**		0.895*					
H'_{OTU}									-0.850*				
L_{phy}	-0.989***					-0.922**	0.838*						
L_{OTU}									-0.950**				

Asco: Ascomycota; Basi: Basidiomycota; Chyt: Chytridiomycota; Crypt: Cryptomycetes; Dacry: Dacrymycetes; Dothi: Dothideomycetes; Glom: Glomeromycota; Sord: Sordariomycetes; Ex Na: Exchangeable Na.

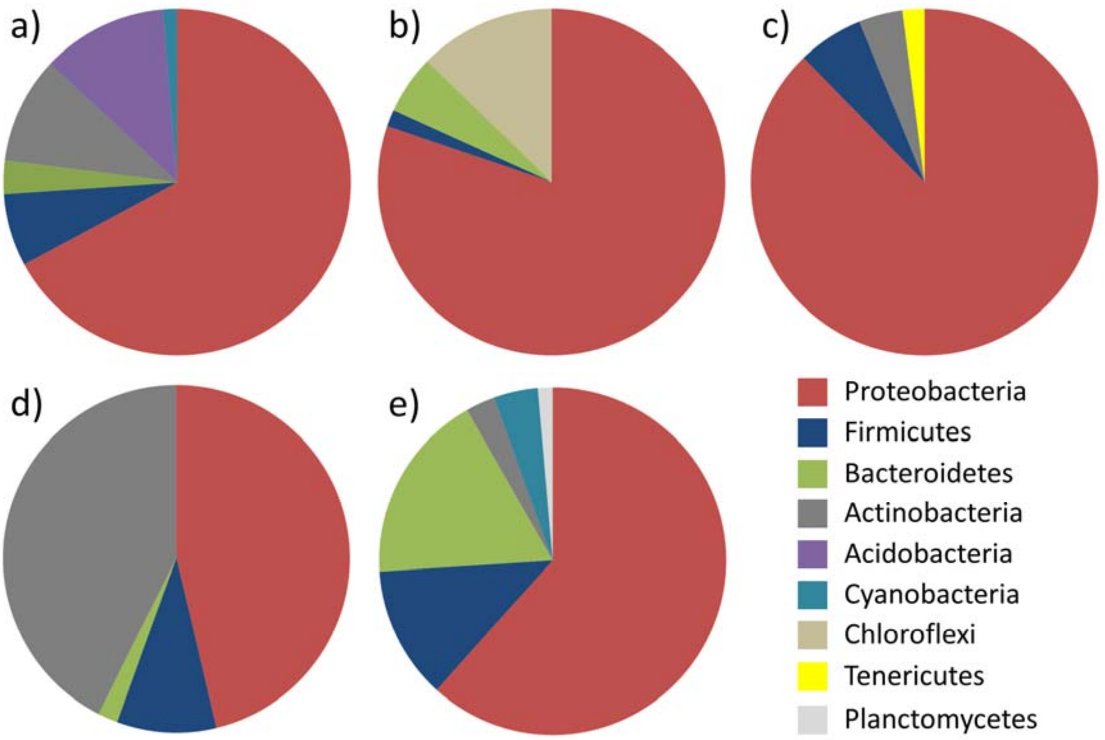

Figure S1. Relative abundance of bacterial phyla as percentage of total sequence reads in current study ((a): bauxite residue, (b) oil sands tailings) and previous studies ((c): chromite ore processing residue; (1), (d) uranium mill tailings; (2), (e) Tirez lagoon; (3). Note that steel slag (4) was not included here as all OTUs belonged to the Proteobacteria, and that Kenyan soda lake (5) sequences were not included as they failed to match any reference sequences during closed reference OTU picking. Phyla representing < 1 % of sequences per environment type are not displayed.

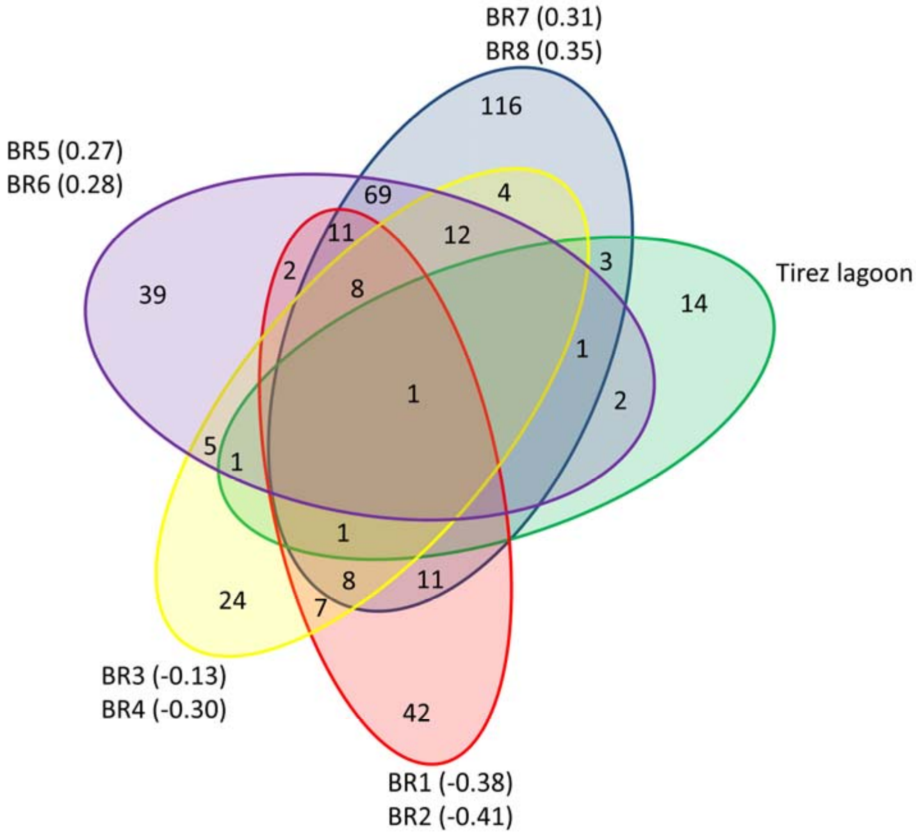

Figure S2. Shared OTUs between bacterial communities in clusters of bauxite residue samples (as identified by UniFrac cluster environments tool) and Tirez lagoon. Each bauxite residue cluster is labelled with the names of samples for which bacterial OTUs were pooled during data analysis, and their PC1 score. Blank regions indicate no shared OTUs.

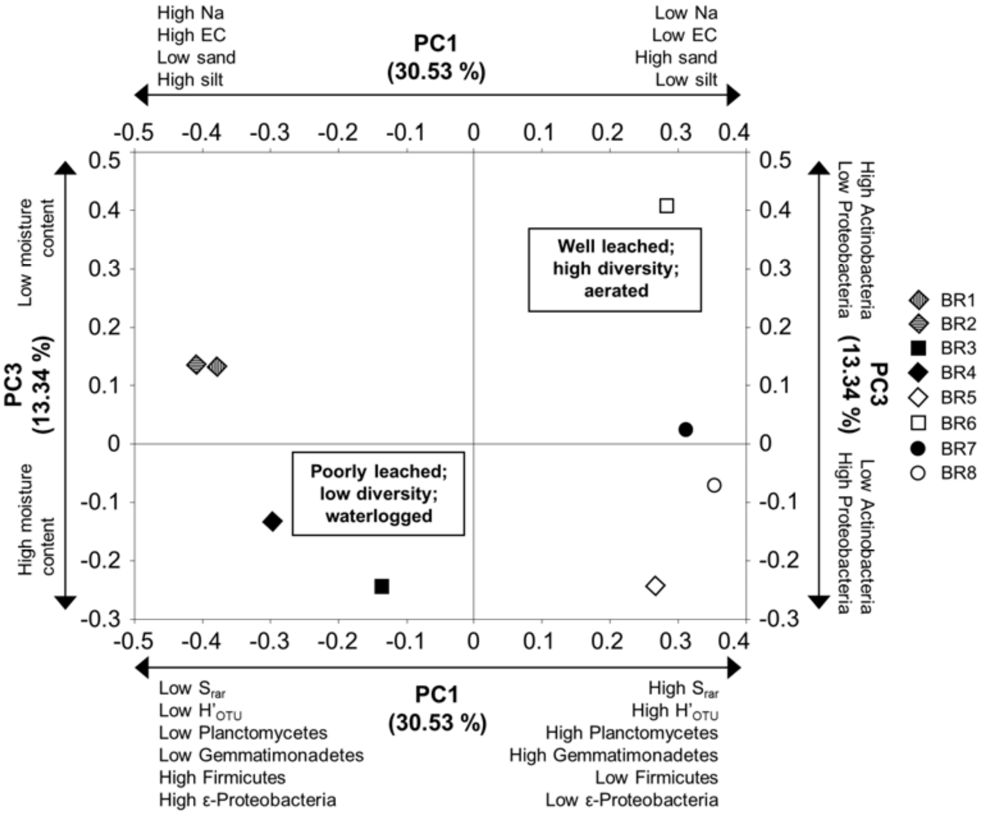

Figure S3. Bacterial community compositions as a function of their loading on principal components 1 (PC1) and 3 (PC3), as determined by UniFrac. Legend refers to samples identified in Table 1. Axis labels refer to principal components (PCs) and the total proportion of variance explained by each component (%). Double headed arrows and labels next to the top X axis and left Y axis identify significant correlations between principal components and environmental variables. Double headed arrows and labels next to the bottom X axis and right Y axis identify significant correlations between principal components and bacterial community composition and diversity variables. Bold text in boxes within graphs identifies inferred environmental characteristics associated with high and low loadings on each combination of principal components.

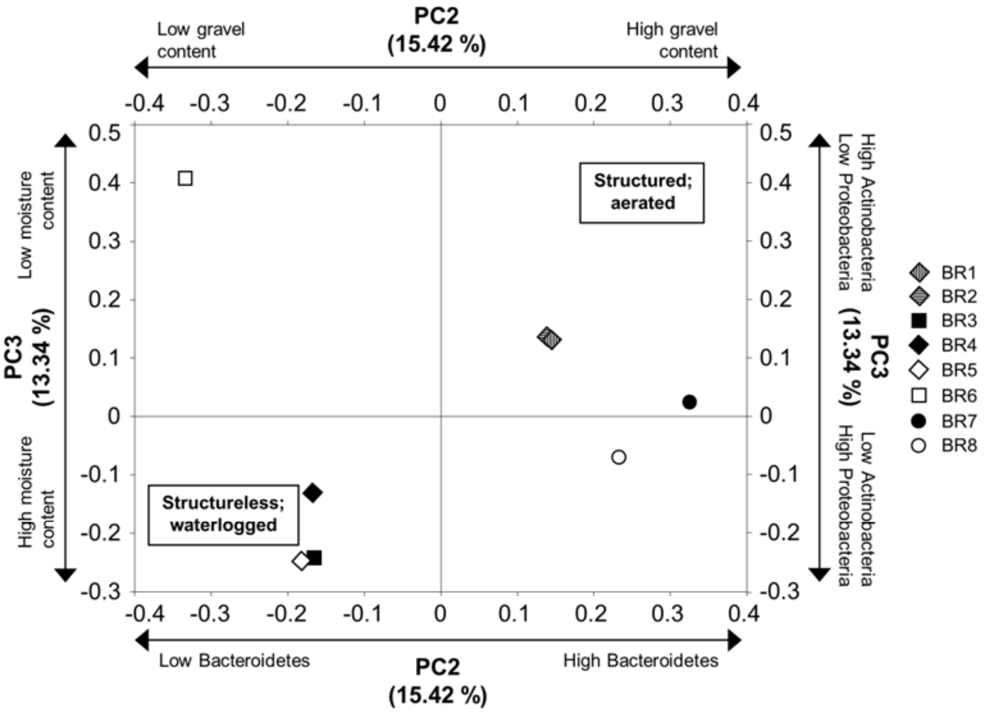

Figure S4. Bacterial community compositions as a function of their loading on principal components 2 (PC2) and 3 (PC3), as determined by UniFrac. Legend refers to samples identified in Table 1. Axis labels refer to principal components (PCs) and the total proportion of variance explained by each component (%). Double headed arrows and labels next to the top X axis and left Y axis identify significant correlations between principal components and environmental variables. Double headed arrows and labels next to the bottom X axis and right Y axis identify significant correlations between principal components and bacterial community composition and diversity variables. Bold text in boxes within graphs identifies inferred environmental characteristics associated with high and low loadings on each combination of principal components.

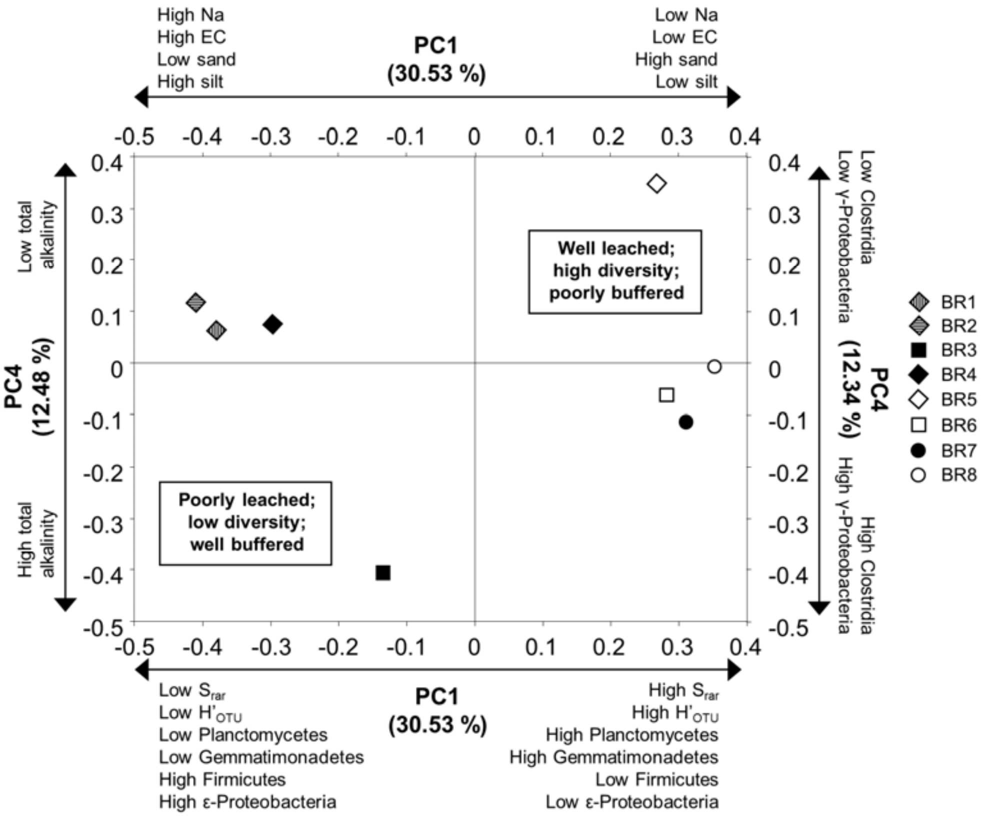

Figure S5. Bacterial community compositions as a function of their loading on principal components 1 (PC1) and 4 (PC4), as determined by UniFrac. Legend refers to samples identified in Table 1. Axis labels refer to principal components (PCs) and the total proportion of variance explained by each component (%). Double headed arrows and labels next to the top X axis and left Y axis identify significant correlations between principal components and environmental variables. Double headed arrows and labels next to the bottom X axis and right Y axis identify significant correlations between principal components and bacterial community composition and diversity variables. Bold text in boxes within graphs identifies inferred environmental characteristics associated with high and low loadings on each combination of principal components.

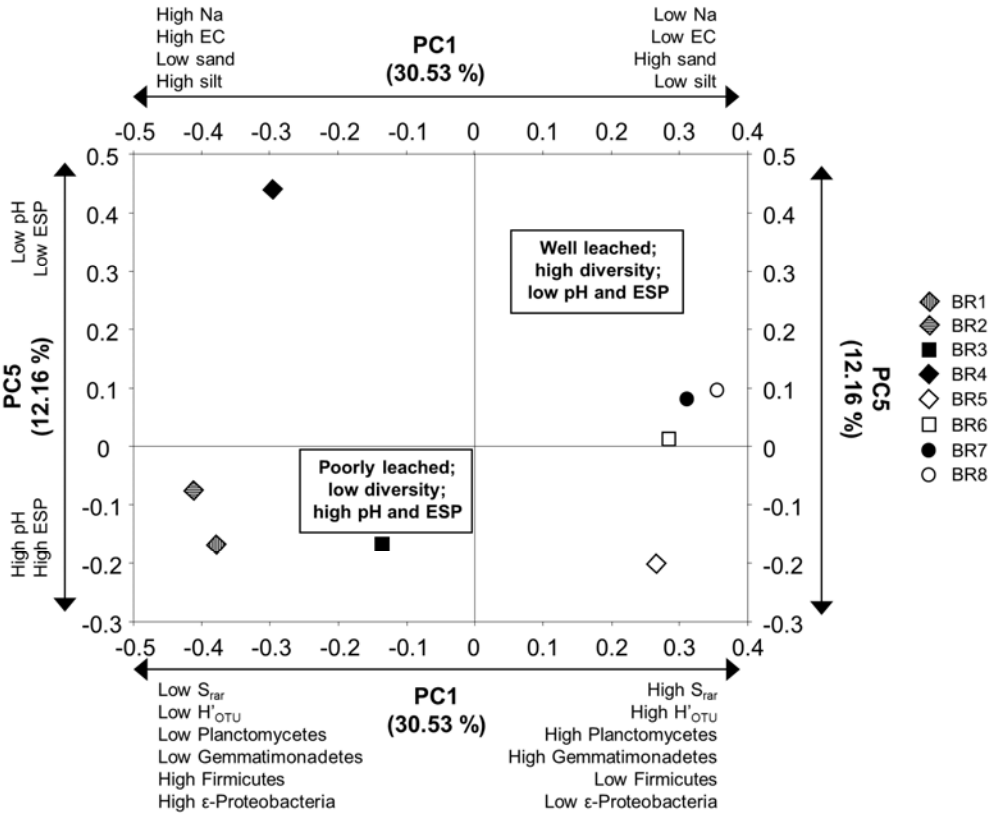

Figure S6. Bacterial community compositions as a function of their loading on principal components 1 (PC1) and 5 (PC5), as determined by UniFrac. Legend refers to samples identified in Table 1. Axis labels refer to principal components (PCs) and the total proportion of variance explained by each component (%). Double headed arrows and labels next to the top X axis and left Y axis identify significant correlations between principal components and environmental variables. Double headed arrows and labels next to the bottom X axis and right Y axis identify significant correlations between principal components and bacterial community composition and diversity variables. Bold text in boxes within graphs identifies inferred environmental characteristics associated with high and low loadings on each combination of principal components.

Table S10. Loadings on principal components for each bacterial community sample analysed, and percentage of total variance explained by each principal component, as determined by the UniFrac tool.

	Principal component						
Sample	1	2	3	4	5	6	7
BR1	-0.378	0.146	0.130	0.061	-0.172	-0.014	0.264
BR2	-0.409	0.140	0.134	0.114	-0.082	-0.038	-0.267
BR3	-0.134	-0.164	-0.245	-0.407	-0.169	-0.018	-0.029
BR4	-0.295	-0.165	-0.133	0.072	0.439	0.042	0.039
BR5	0.267	-0.181	-0.246	0.345	-0.204	0.084	-0.005
BR6	0.283	-0.333	0.408	-0.064	0.012	-0.016	-0.002
BR7	0.313	0.325	0.024	-0.113	0.080	0.339	-0.012
BR8	0.354	0.233	-0.071	-0.008	0.096	-0.381	0.011
Total variation explained (%)	30.53	15.42	13.34	12.48	12.16	10.49	5.56

Table S11. Loadings on principal components for each fungal community sample analysed, and percentage of total variance explained by each principal component, as determined by the UniFrac tool.

	Principal component				
Sample	1	2	3	4	5
BR1	-	-	-	-	-
BR2	0.102	-0.341	-0.366	-0.008	0.010
BR3	0.665	0.130	0.116	0.000	0.016
BR4	-	-	-	-	-
BR5	-0.204	0.217	-0.059	-0.355	-0.008
BR6	-0.195	0.309	-0.154	0.273	-0.078
BR7	-0.149	-0.228	0.278	0.015	-0.232
BR8	-0.218	-0.087	0.185	0.075	0.292
Total variation explained (%)	38.21	21.25	18.17	13.11	9.25

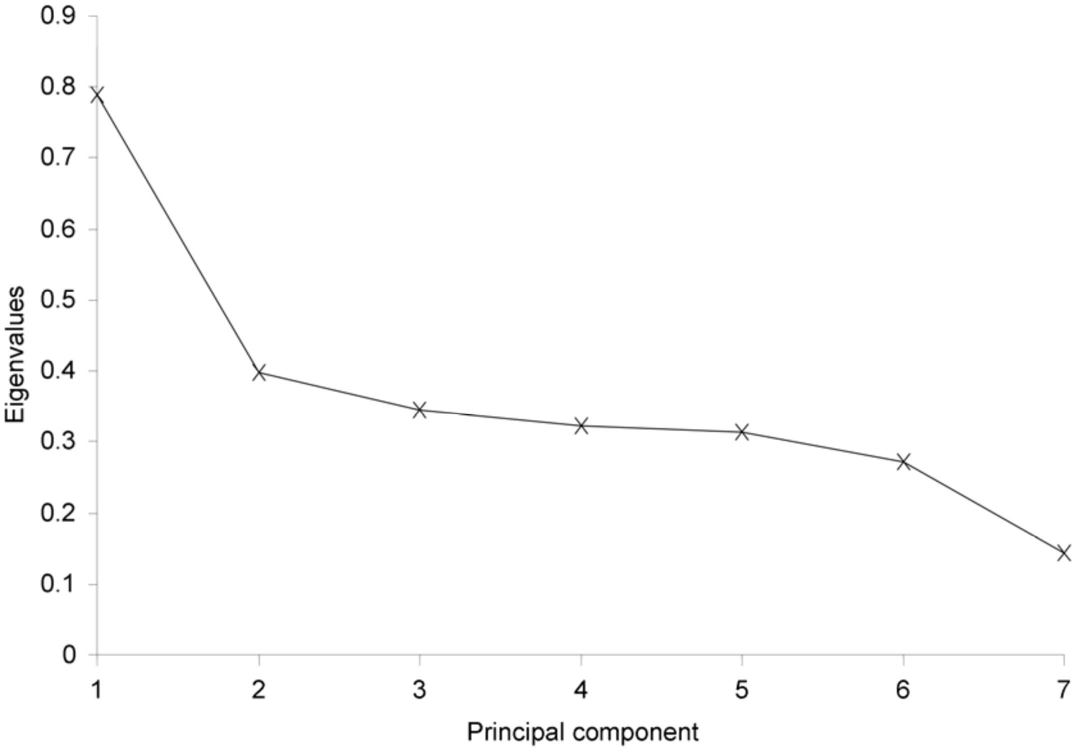

Figure S7. Scree plot illustrating eigenvalues for each principal component identified in the analysis of bacterial communities within bauxite residue samples with the UniFrac tool.

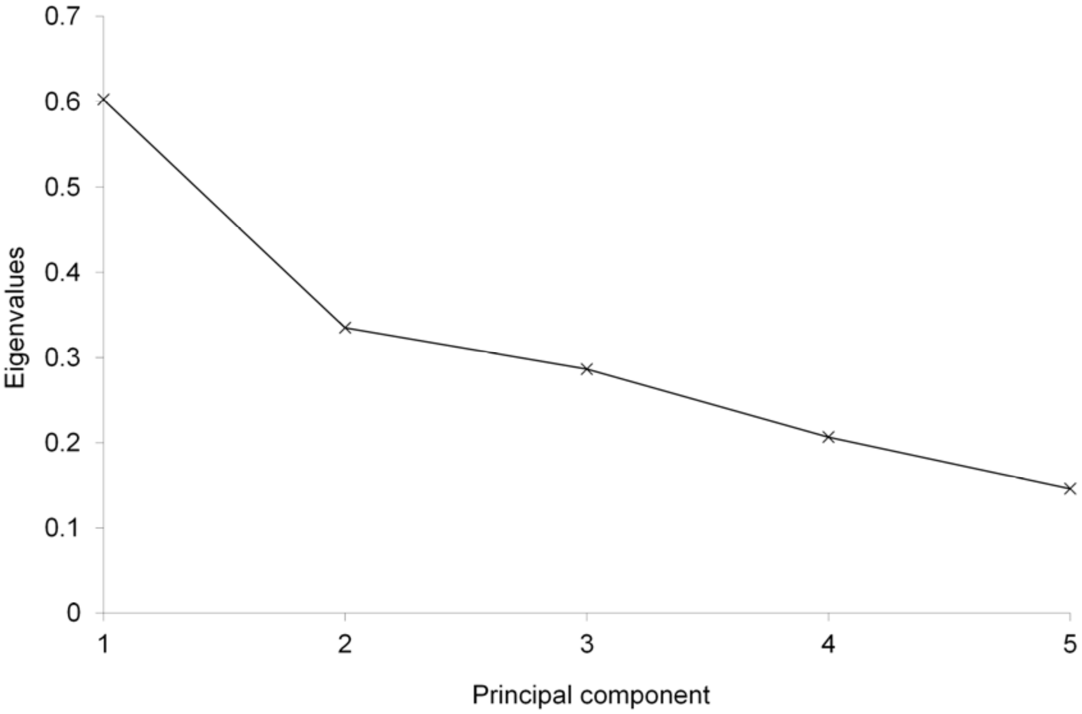

Figure S8. Scree plot illustrating eigenvalues for each principal component identified in the analysis of fungal communities within bauxite residue samples with the UniFrac tool.

References

1. **Brito EMS, Piñón-Castillo HA, Guyoneaud R, Caretta CA, Gutiérrez-Corona JF, Duran R, Reyna-López GE, Nevárez-Moorillón GV, Fahy A, Goñi-Urriza M.** 2013. Bacterial biodiversity from anthropogenic extreme environments: a hyper-alkaline and hyper-saline industrial residue contaminated by chromium and iron. Appl Microbiol Biotechnol **97**:369-378.
2. **Bondici VF, Lawrence JR, Khan NH, Hill JE, Yergeau E, Wolfaardt GM, Warner J, Korber DR.** 2013. Microbial communities in low permeability, high pH uranium mine tailings: characterisation and potential effects. J Appl Microbiol **114**:1671-1686.
3. **Montoya L, Vizioli C, Rodríguez N, Rastoll MJ, Amils R, Marin I.** 2013. Microbial community composition of Tirez lagoon (Spain), a highly sulphated athalassohaline environment. Aquat Biosyst **9**:19.
4. **Roadcap GS, Sanford RA, Jin Q, Pardinas JR, Bethke CM.** 2006. Extremely alkaline (pH>12) groundwater hosts diverse microbial community. Groundwater **44**:511-517.
5. **Rees HC, Grant WD, Jones BE, Heaphy S.** 2004. Diversity of Kenyan soda lake alkaliphiles assessed by molecular methods. Extremophiles **8**:63-71.