Bacterial diversity in the waterholes of the Kruger National Park: an eDNA metabarcoding approach
Supplementary Materials

M.J. Farrell, D. Govender,
M. Hajibabaei, M. van der Bank, T. J. Davies

Table 1: Water quality measurements

| Sample code	Site	Date	Temp (°C)	mS/cm	DO (%)	DO (mg/L)	pH
DLP_8	DLP	July 10	15.27	3.11	83.37	39.67	9.16
GIR_1	GIR	June 24	18.58	1.95	50.83	42.00	9.27
GIR_2	GIR	July 1	21.85	1.80	74.47	41.00	9.24
GIR_3	GIR	July 8	20.72	1.90	88.47	39.00	9.35
HOY_2	HOY	June 22	17.59	3.18	14.53	43.43	8.14
HOY_3	HOY	June 29	17.84	3.01	42.53	40.00	8.25
HOY_4	HOY	July 6	16.83	2.96	39.27	35.90	8.39
IMB_2	IMB	June 22	15.17	2.46	74.80	46.77	8.19
IMB_3	IMB	June 29	16.07	2.43	45.23	40.67	8.16
IMB_4	IMB	July 6	15.56	2.46	35.30	35.57	8.13
KWA_5	KWA	June 19	19.20	1.80	154.90	55.30	9.92
KWA_6	KWA	June 26	16.50	1.75	111.93	45.10	9.51
NGO_2	NGO	June 24	14.97	0.48	111.23	44.10	9.44
NGO_3	NGO	July 1	17.26	0.48	108.43	40.00	9.45
NGO_4	NGO	July 8	18.22	0.51	94.50	40.33	9.21
NHL_2	NHL	June 22	17.80	1.99	118.23	50.20	8.45
NHL_3	NHL	June 29	25.52	1.97	134.97	49.87	8.08
NHL_4	NHL	July 6	22.38	2.03	125.73	40.93	8.23
NWA_2	NWA	June 26	16.14	0.90	106.67	44.77	9.77
NWA_3	NWA	June 29	24.14	0.82	199.27	53.60	9.90
NWA_4	NWA	June 30	18.91	0.90	124.63	44.43	9.66
NWA_5	NWA	July 1	23.40	0.93	180.93	49.20	9.75
NWA_6	NWA	July 2	18.65	0.91	114.90	41.00	9.66
NWA_7	NWA	July 3	17.90	0.91	104.30	40.00	9.73
NWA_8	NWA	July 10	18.06	0.87	68.53	37.27	9.24
NYA_2	NYA	June 24	14.92	0.52	54.90	39.67	8.64
NYA_3	NYA	July 1	17.51	0.54	70.27	38.00	8.64
NYA_4	NYA	July 8	18.30	0.55	76.93	37.63	8.82
WIT_2	WIT	June 24	15.31	0.58	173.60	48.87	9.40
WIT_3	WIT	July 1	18.94	0.58	139.73	43.73	9.11
WIT_4	WIT	July 8	18.89	0.69	63.37	36.23	8.54
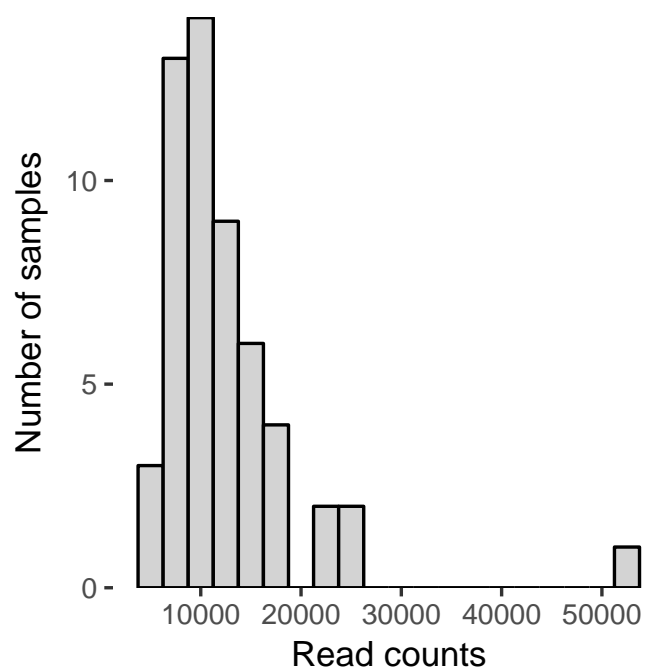
Figure 1: Histogram of retained sequencing depth across samples.
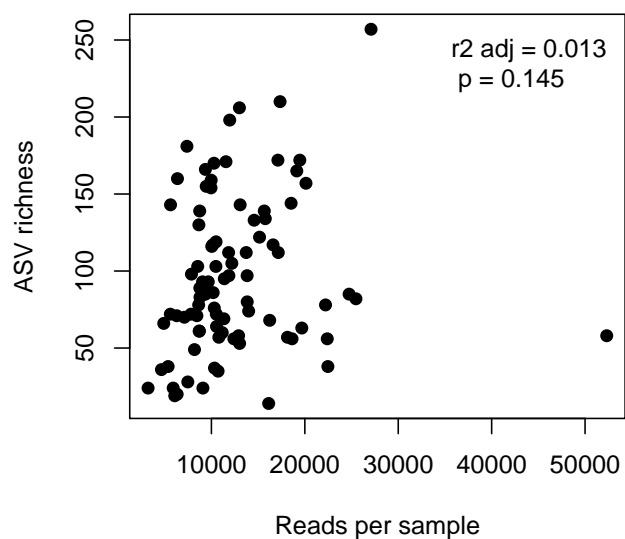

Figure 2: Scatterplot of reads and ASV richness per sample.

Table 2: Read counts tracked through the DADA2 pipeline

Sample code	Input	Filtered	Denoised	Merged	No Chimeras	Final Read Count	ASV Richness
BLANK_2	33813	24642	24642	22763	22468	22468	38
BLANK_3	20304	17219	17219	16166	16130	16130	14
DLP_8A	27355	26994	26994	22875	15843	15767	136
DLP_8A_S	31978	30863	30863	25222	19664	19461	179

DLP_8A_XS	38337	36972	36972	31551	27325	27087	274
DLP_8B	48282	40544	40544	34017	22401	22401	57
DLP_8B_S	28748	27713	27713	22625	17294	17119	176
DLP_8B_XS	27466	26385	26385	22277	19193	19138	168
GIR_1A	17079	16649	16649	12463	8752	8658	146
GIR_1B	15040	14314	14314	10702	7928	7865	102
GIR_2A	19825	18985	18985	14444	10003	9974	179
GIR_2B	25000	23992	23992	18315	11595	11558	199
GIR_3A	19792	19024	19024	15107	10745	10479	111
GIR_3B	23759	22888	22888	18333	12489	12184	116
HOY_2A	33258	32209	32209	15900	13831	13831	83
HOY_2B	67589	61567	61567	55496	52321	52321	58
HOY_3A	13966	13303	13303	6637	5364	5364	38
HOY_3A_S	31543	29962	29962	19154	15199	15152	128
HOY_3A_XS	34251	32392	32392	20304	16688	16599	130
HOY_3B	46882	40568	40568	31787	25515	25470	87
HOY_3B_S	14641	13727	13727	7053	5936	5908	24
HOY_3B_XS	39784	37517	37517	23198	18706	18531	168
HOY_4A	21589	19517	19517	13064	10283	10278	192
HOY_4B	26018	23698	23698	15860	11959	11959	223
IMB_2A	25930	22856	22856	15788	13076	13076	149
IMB_2B	32851	29178	29178	21347	17355	17349	227
IMB_3A	28191	25483	25483	16789	12998	12998	231
IMB_3B	21749	18798	18798	11600	9444	9444	177
IMB_4A	33674	30905	30905	19937	15659	15659	151
IMB_4B	32822	30115	30115	18815	14584	14578	142
KWA_5A	46221	39502	39502	29959	25293	24740	88
KWA_5B	44409	38312	38312	28420	22721	22216	80
KWA_6A	15998	15708	15708	12820	10654	9660	94
KWA_6A_S	21283	20439	20439	17043	15071	13822	98
KWA_6A_XS	31733	30158	30158	25821	24649	20123	162
KWA_6B	29050	24618	24618	17083	15058	12912	60
KWA_6B_S	18791	18050	18050	15145	13585	11870	100
KWA_6B_XS	44397	42687	42687	24822	24362	19676	63
NGO_2A	23500	22834	22834	18953	15565	10494	135
NGO_2B	34501	29490	29490	23297	20611	16238	72
NGO_3A	23725	23459	23459	19390	13459	8635	83
NGO_3A_S	20459	19686	19686	15621	12597	10548	67
NGO_3A_XS	25913	25002	25002	14069	13273	7472	28
NGO_3B	19885	16860	16860	11958	9884	6077	19
NGO_3B_S	23204	22345	22345	17909	14328	10529	73
NGO_3B_XS	23313	22525	22525	13626	12918	9092	24
NGO_4A	22054	21360	21360	16498	12455	10794	59
NGO_4B	22608	21868	21868	17077	13071	11159	60
NHL_2A	30627	29976	29976	26084	17626	17156	134
NHL_2B	15758	15015	15015	11797	8362	8185	53
NHL_3A	20174	18424	18424	13628	10730	9355	181
NHL_3B	17707	15991	15991	11497	9145	8746	142
NHL_4A	16675	15894	15894	12917	8961	8700	68
NHL_4B	20089	19130	19130	15677	10697	10307	83
NWA_2A	26097	22271	22271	16953	12983	12405	57

NWA_2A_S	19149	18326	18326	15107	12871	11840	116
NWA_2A_XS	20071	19273	19273	10180	9890	6330	20
NWA_2B	20898	17634	17634	13251	10543	10330	37
NWA_2B_S	15648	14878	14878	12009	9819	9052	104
NWA_2B_XS	27808	26838	26838	13564	12591	10713	38
NWA_3A	10226	9735	9735	7264	6029	4895	67
NWA_3B	28430	26973	26973	21273	15416	13726	119
NWA_4A	17654	17123	17123	14118	11196	10140	124
NWA_4B	15907	15184	15184	12240	9456	8760	93
NWA_5A	14896	14275	14275	11127	9207	8777	91
NWA_5B	17054	16316	16316	13064	10570	10189	89
NWA_6A	19653	18822	18822	14894	12020	11380	98
NWA_6B	15184	14591	14591	11804	9839	9355	85
NWA_7A	13767	13177	13177	10932	9017	8445	72
NWA_7B	16831	16102	16102	12835	9883	9231	87
NWA_8A	15266	14481	14481	10673	8879	8531	109
NWA_8B	17882	16958	16958	12503	10249	10038	127
NYA_2A	33395	28245	28245	21413	18148	18143	60
NYA_2B	36205	30519	30519	23387	18678	18631	57
NYA_3A	29675	25050	25050	16505	13118	13019	55
NYA_3B	30193	25942	25942	17723	14146	13985	76
NYA_4A	17561	16928	16928	12689	8760	7377	241
NYA_4A_S	13136	12431	12431	9507	7404	5624	157
NYA_4A_XS	13283	12762	12762	5737	4120	3220	24
NYA_4B	24397	20885	20885	14109	9512	7813	74
NYA_4B_S	15610	14774	14774	11127	8646	6365	169
NYA_4B_XS	16154	15502	15502	8077	6182	4653	36
WIT_2A	14307	13757	13757	9617	7282	7100	78
WIT_2B	13757	13277	13277	8501	6283	6283	80
WIT_3A	25714	25291	25291	16183	10271	9958	201
WIT_3B	14067	13479	13479	7899	5799	5600	76
WIT_4A	18581	17947	17947	13911	8764	8701	66
WIT_4B	22216	21441	21441	16481	11437	11333	78
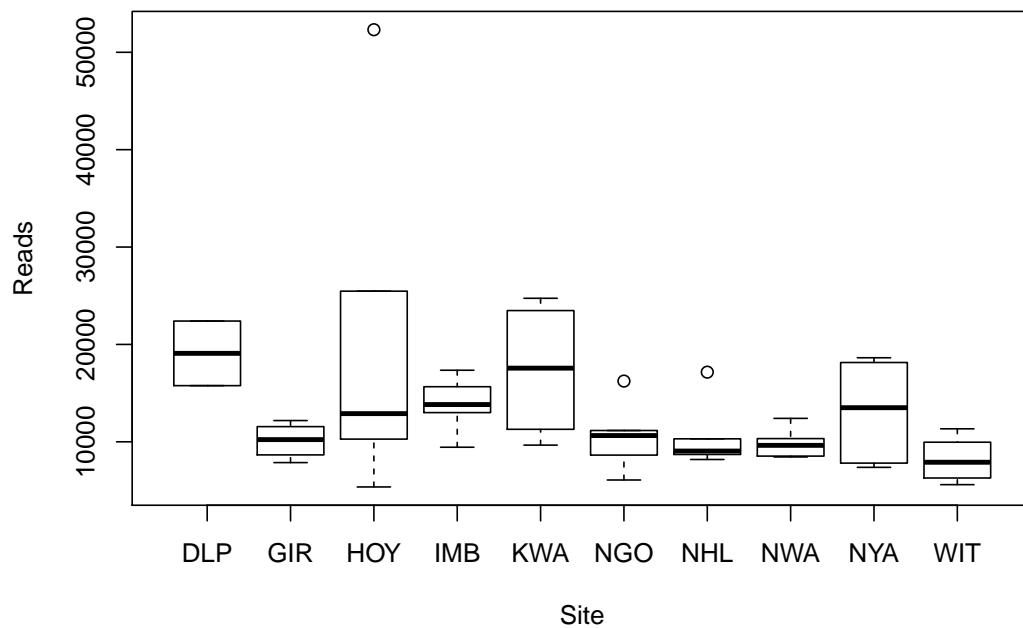
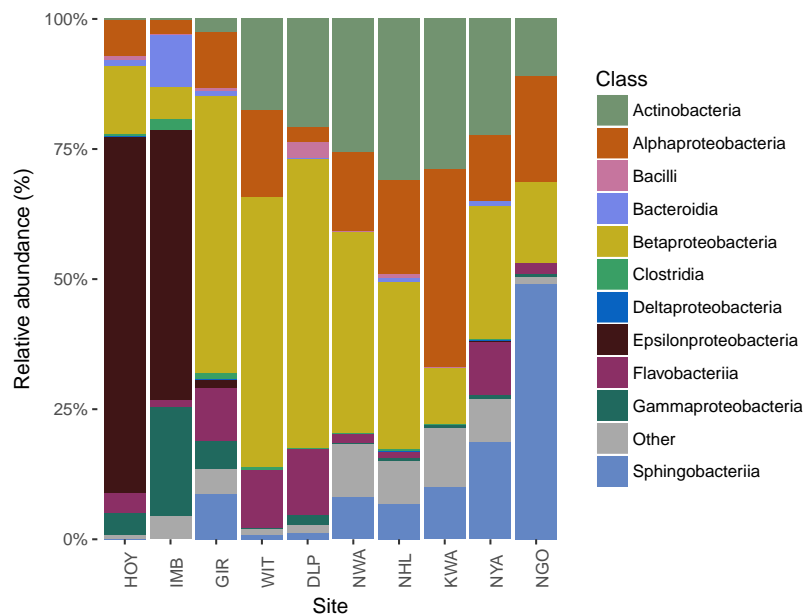Figure 3: Box and whisker plots of reads among the core samples, grouped by site.

Figure 4: Relative abundances of bacterial classes across sites. Sites are ordered by relative abundance of phylum Proteobacteria.
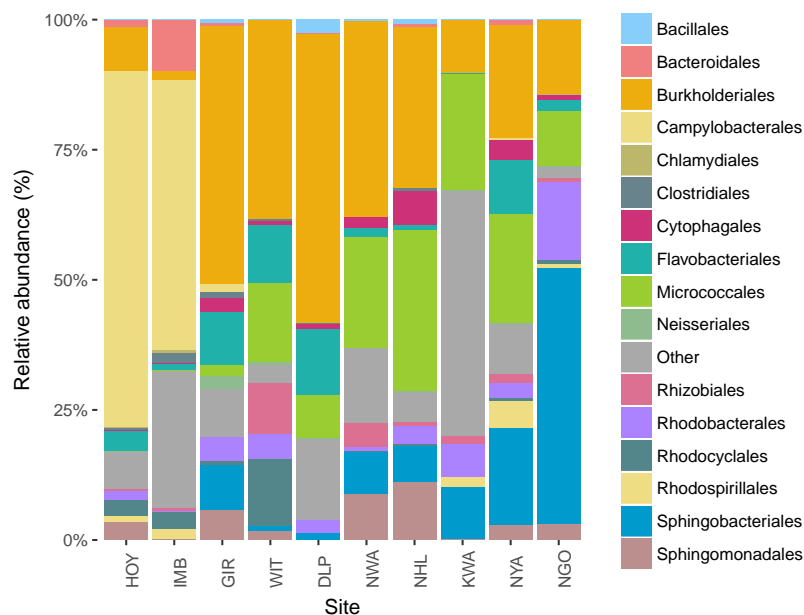
Figure 5: Relative abundances of bacterial orders across sites. Sites are ordered by relative abundance of phylum Proteobacteria.
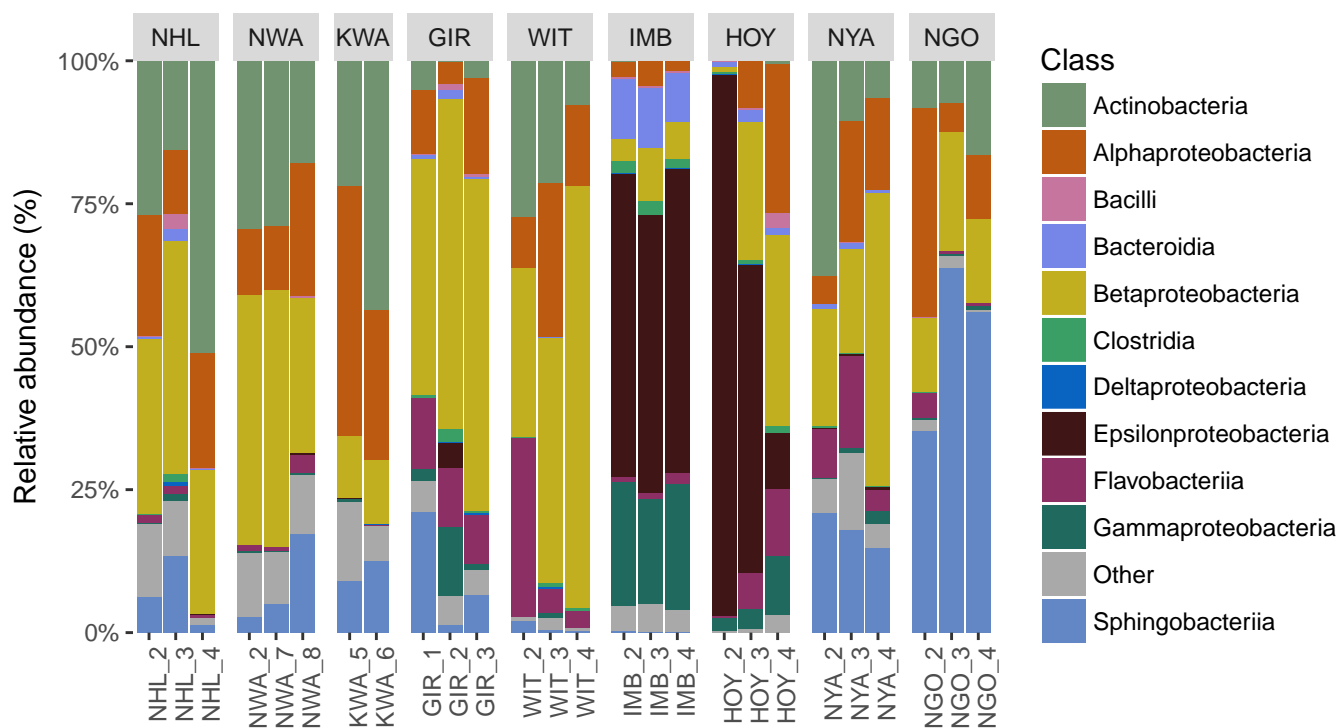

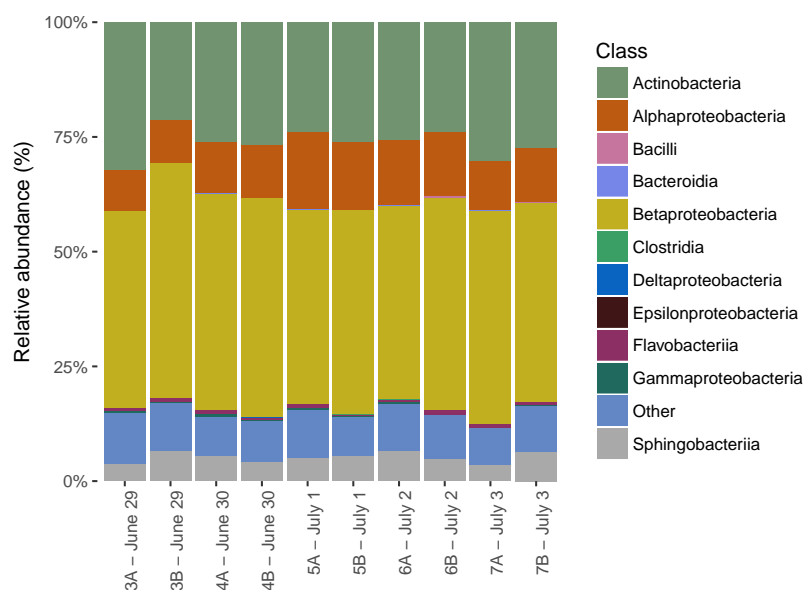Figure 6: Relative abundances of bacterial classes across weekly samples.

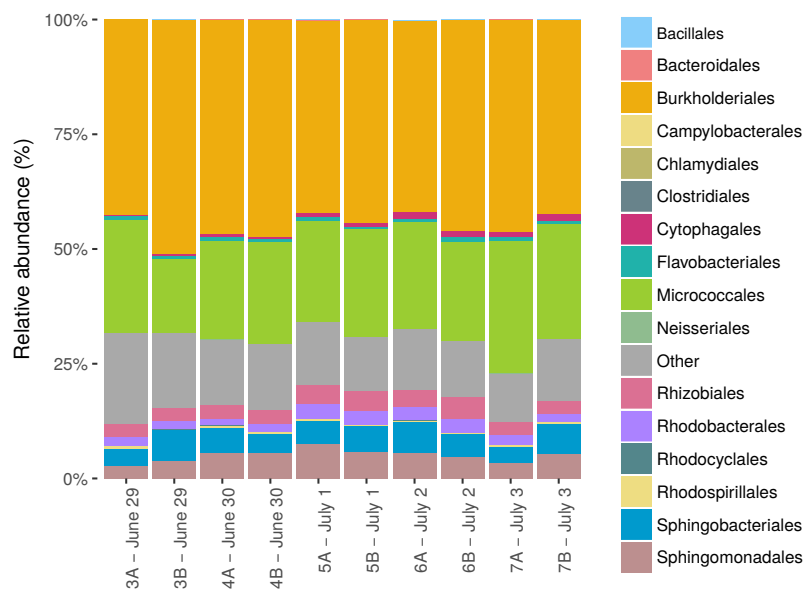Figure 7: Relative abundances of bacterial classes across five days at a single site (NWA).

Figure 8: Relative abundances of bacterial orders across five days at a single site (NWA).
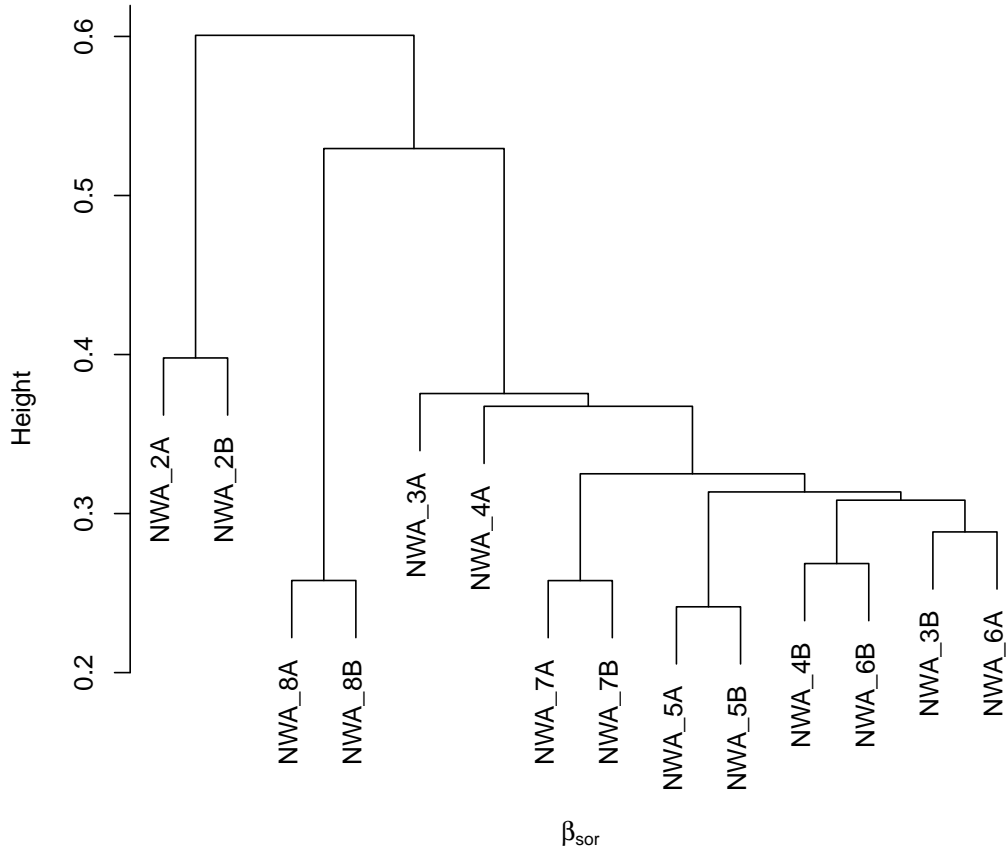
Figure 9: Hierarchical clustering of 150 mL NWA samples based on pairwise Sorensen's beta diversity.

(a) Bray-Curtis NMDS

(b) Weighted UniFrac NMDS
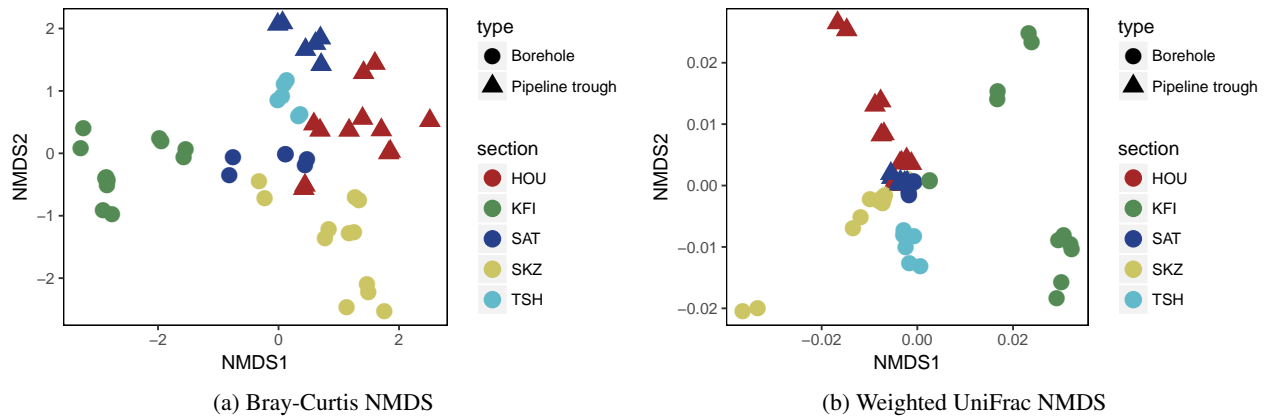
Figure 10: NMDS plots of a) Bray-Curtis and b) abundance-weighted UniFrac distances. Colours represent section, shapes represent waterhole type.

(a) Bray-Curtis NMDS

(b) Weighted UniFrac NMDS
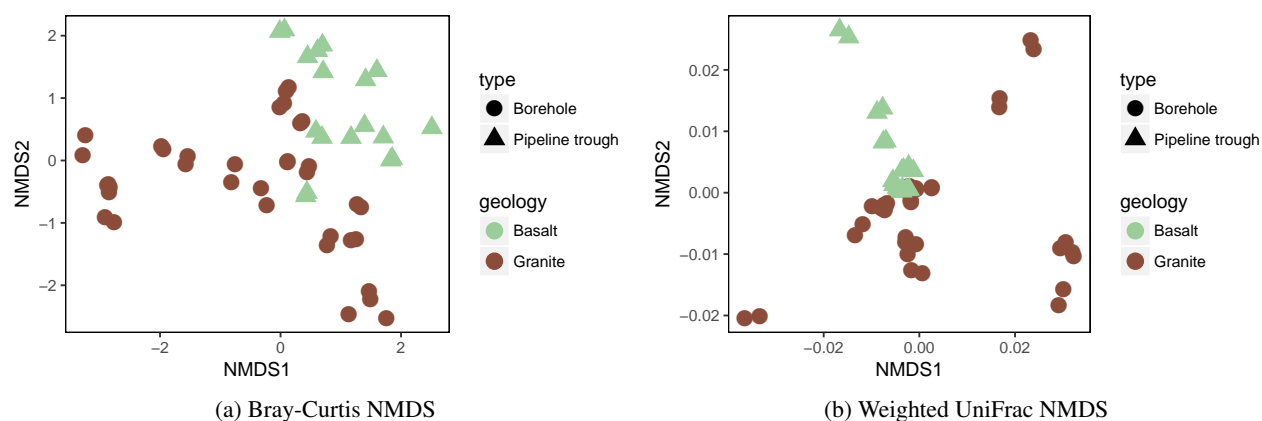
Figure 11: NMDS plots of a) Bray-Curtis and b) abundance-weighted UniFrac distances. Colours represent subsurface geology, shapes represent waterhole type.

(a) Bray-Curtis NMDS

(b) Weighted UniFrac NMDS
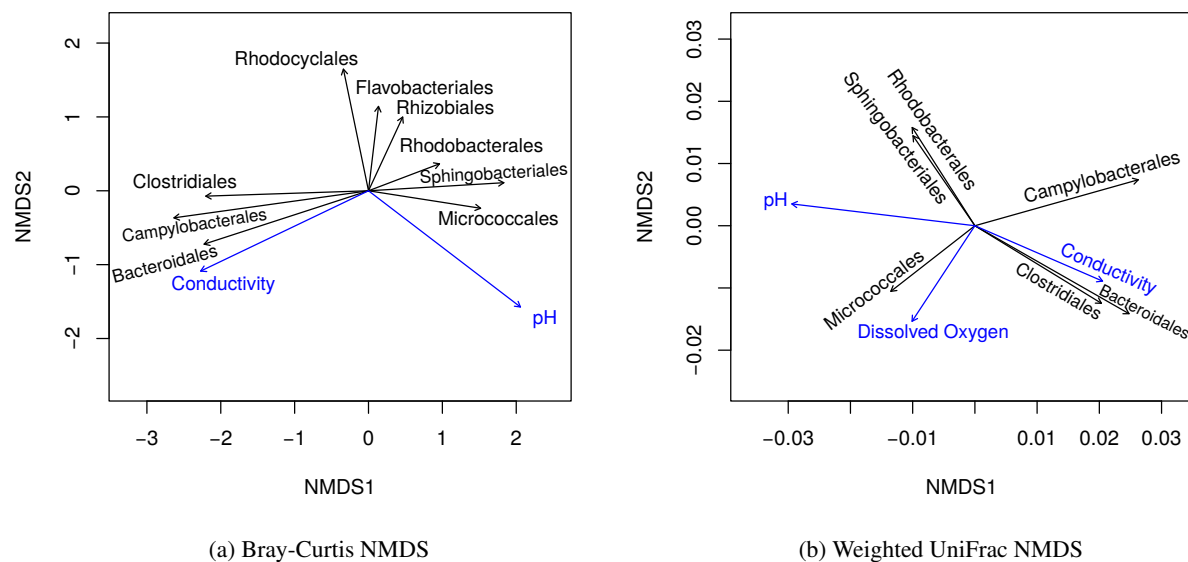
Figure 12: Nonmetric multidimensional scaling (NMDS) ordination of variation in bacterial community structure across 54 samples based on a) Bray-Curtis and b) abundance-weighted UniFrac distances. Arrows indicate the direction of significant ($p < 0.05$) correlations among variables and the NMDS axes, with arrow length indicating the strength of the correlation. Blue arrows indicate environmental variables, while black arrows indicate relative abundances of sequences from different microbial orders. The ordination axes explain 96.9% (a) and 98.1% (b) of the variance in the dissimilarities (Fig. SM 17).
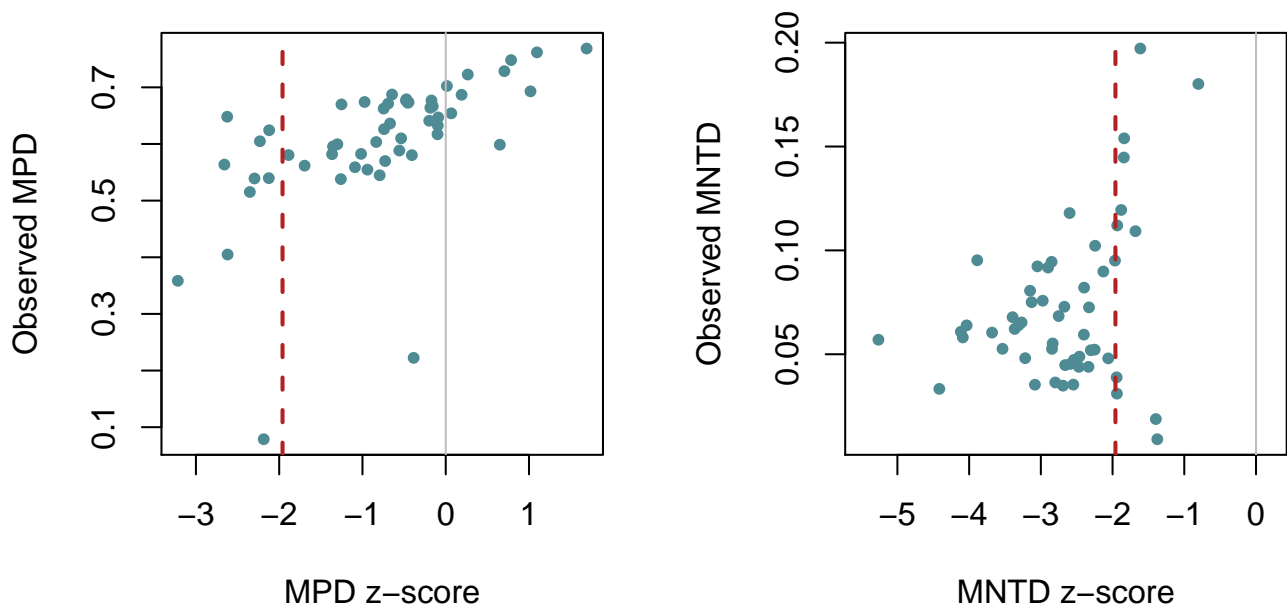
Figure 13: Phylogenetic community structure of bacterial ASVs across samples based on A) mean phylogenetic pairwise distance (MPD) and B) mean nearest taxon distance (MNTD). Each plot depicts the relationship between observed values (y-axis), and the standardized effect size (z-score) for each community following 999 permutaions using the "richness" null model. Gray lines depict z-scores of 0. Red dashed lines represent critical values for a two-tailed z-test (α=0.05), with points lesser than these values indicating significant phylogenetic clustering.
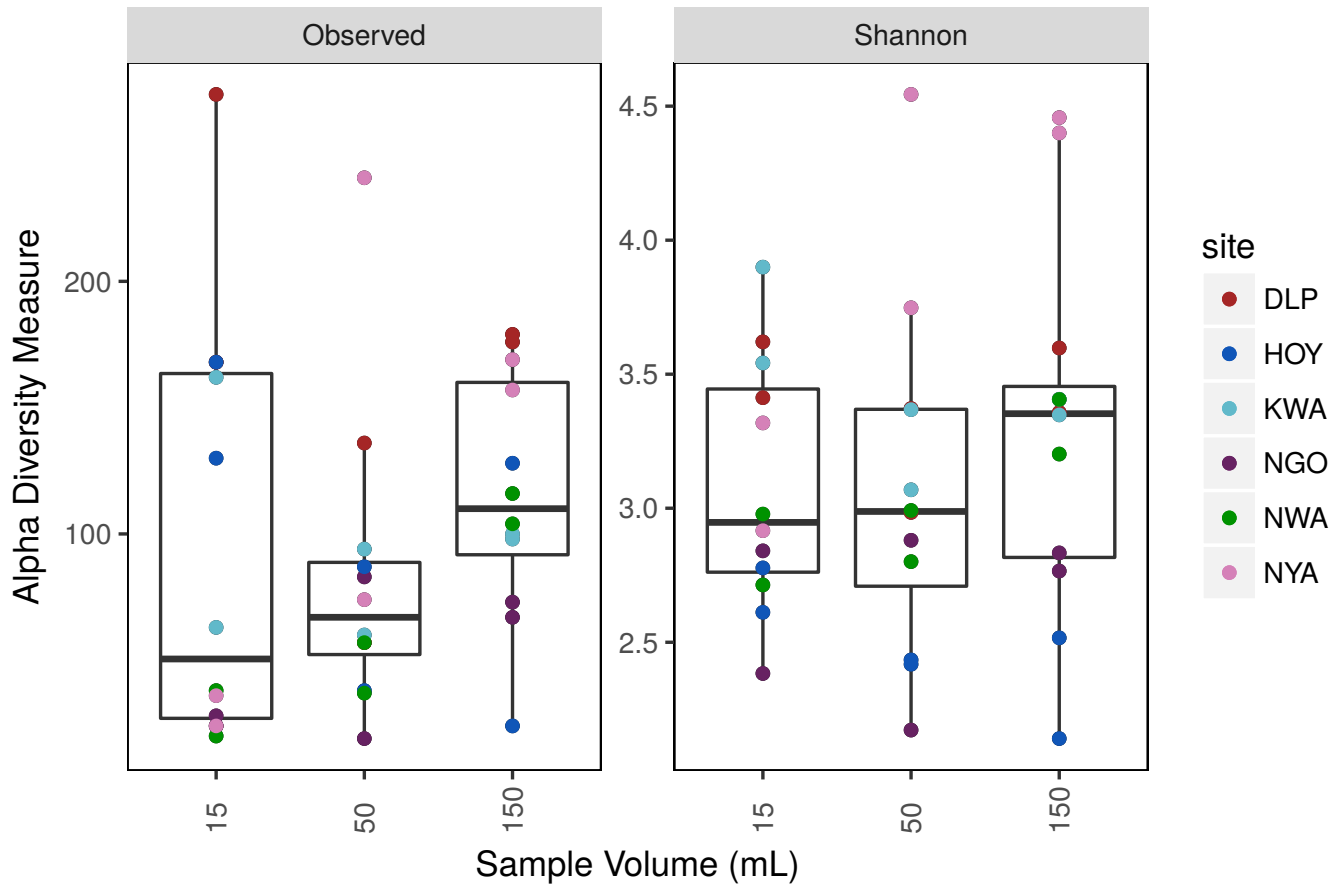
Figure 14: Alpha diversity as measured by observed number of ASVs and Shannon diversity for samples in which 150 mL, 50 mL and 15 mL volumes were filtered. Colours represent sites.
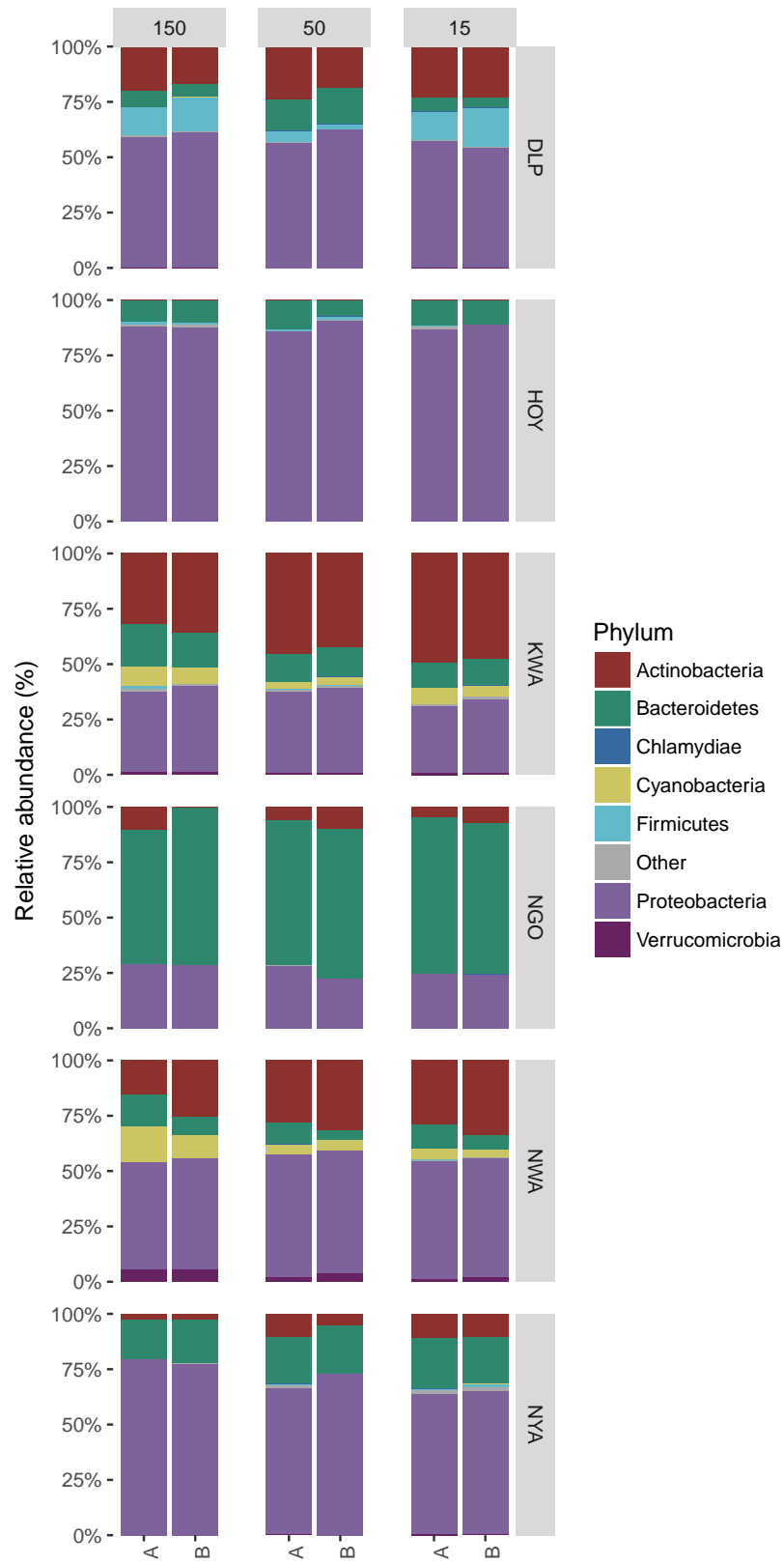
Figure 15: Relative abundances of bacterial phyla across different sample volumes.
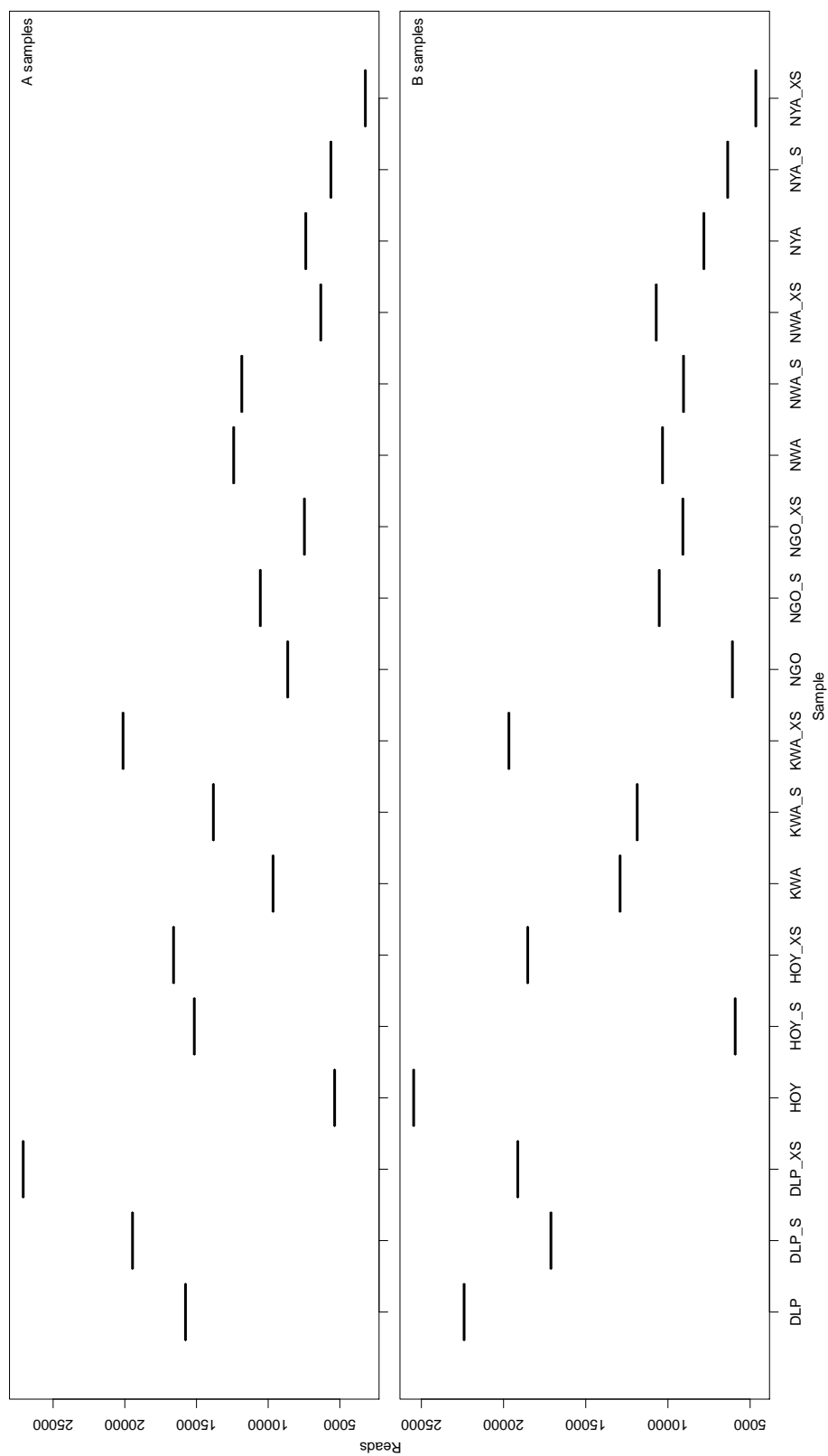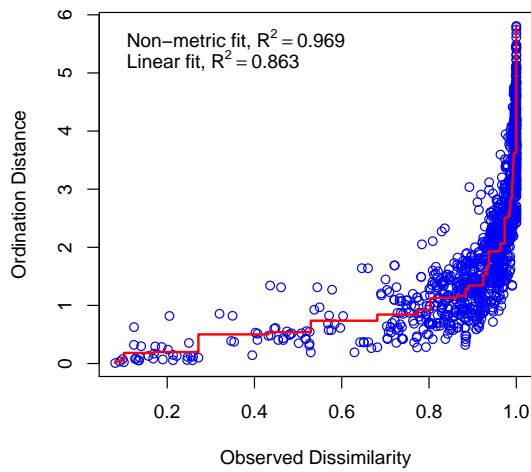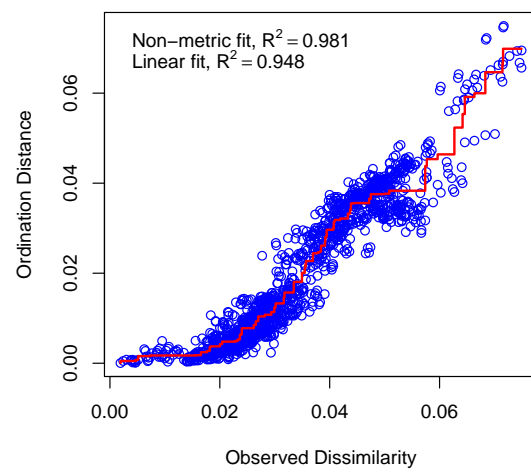
Figure 16: Reads for samples in which different volumes were filtered.

13

(a) Bray-Curtis NMDS

(b) Weighted UniFrac NMDS

Figure 17: Stressplots for NMDS plots using a) Bray-Curtis and b) abundance-weighted UniFrac distances