

Supplementary table 1: The pre- and post-assembly read numbers associated with each sample. The number of contigs exceeding 1000bp, for both the *CLC Genomics Workbench* and *metaSPAdes* assemblies, are listed. The datasets are linked to sample specific NCBI Biosample accession numbers.

Sample	Number of reads before trimming	Number of reads after trimming	Number of contigs, CLC Genomics Workbench (>1000bp)	Number of contigs, metaSPAdes Workbench (>1000bp)	NCBI Biosample accession number
19-3001	14,488,694	10,933,266	764	747	SRR11603223
19-3004	1,759,072	1,483,805	138	217	SRR11603222
19-3005	2,333,498	1,949,519	215	418	SRR11603212
19-3006	1,327,624	1,088,164	132	233	SRR11603211
19-3007	1,309,512	1,042,998	81	110	SRR11603210
19-3008	1,648,818	1,464,956	111	183	SRR11603209
19-3009	1,396,470	1,230,058	157	295	SRR11603208
19-3010	1,130,528	972,354	148	229	SRR11603207
19-3013	1,057,662	837,497	84	142	SRR11603206
19-3017	1,488,540	1,174,523	111	209	SRR11603205
19-3021	1,019,462	780,404	76	109	SRR11603220
19-3035	2,877,330	2,744,446	558	1661	SRR11603215
19-3019	1,282,544	1,123,948	184	295	SRR11603221
19-3026	1,298,852	1,148,624	142	243	SRR11603219
19-3027	2,284,568	2,075,296	307	533	SRR11603218
19-3040	3,502,290	3,255,733	458	780	SRR11603214
19-3044	2,940,068	2,741,705	514	943	SRR11603213
19-3029	3,305,238	2,930,254	224	455	SRR11603217
19-3032	3,713,604	3,501,749	668	1250	SRR11603216
Total	50,164,374	42,479,299			