

1 **SUPPLEMENTARY MATERIAL**

2 for

3 Next-generation sequencing of microbial communities in the Athabasca River
4 and its tributaries in relation to oil sands mining activities

5
6
7 Etienne Yergeau¹, John R. Lawrence², Sylvie Sanschagrin¹, Marley J. Waiser², Darren R.
8 Korber³ and Charles W. Greer¹

9
10 ¹National Research Council Canada, Montreal, Quebec, Canada

11 ²Environment Canada, Saskatoon, Saskatchewan, Canada

12 ³University of Saskatchewan, Saskatoon, Saskatchewan, Canada

Figure S1. Principal coordinate analyses (PCoA) based on Unifrac distances for 454 and Ion Torrent bacterial datasets. AR, Athabasca reference site; NSM, Northland sawmill; DSU, downstream Suncor; MSK, Muskeg River; AER, mouth of Ells River; EU, EM, EL: upper, mid and lower Ells River; FU, FM, FL: upper, mid and lower Firebag Creek; SU, SM, SL: upper, mid and lower Steepbank Creek; SY1A, SY1B, Syncrude tailings ponds; SU3: Suncor tailings pond.

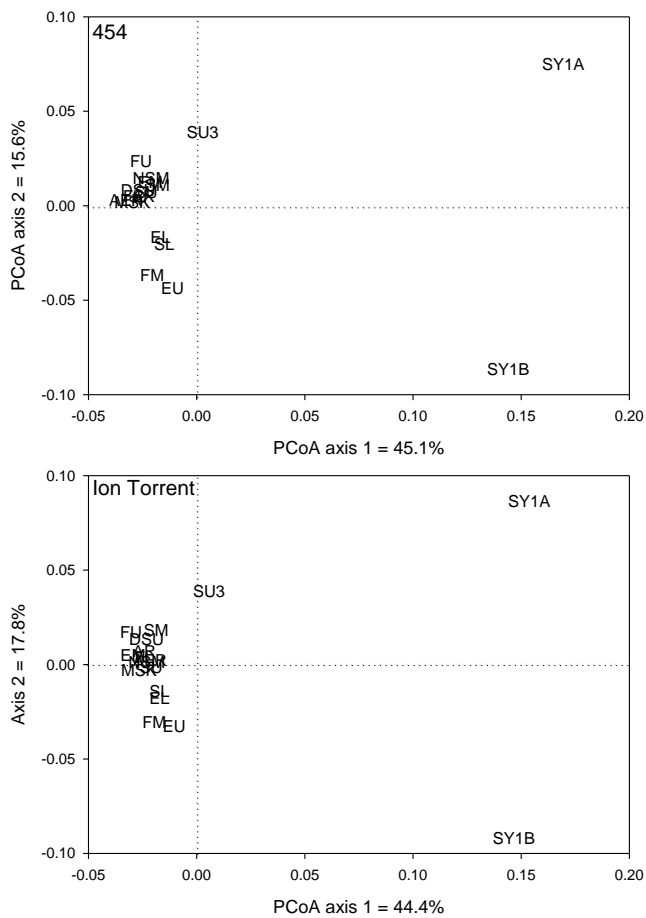


Figure S2. Principal coordinate analyses (PCoA) based on Unifrac distances for Ion Torrent bacterial datasets. Refer to the legend of Figure S1 for sample abbreviations.

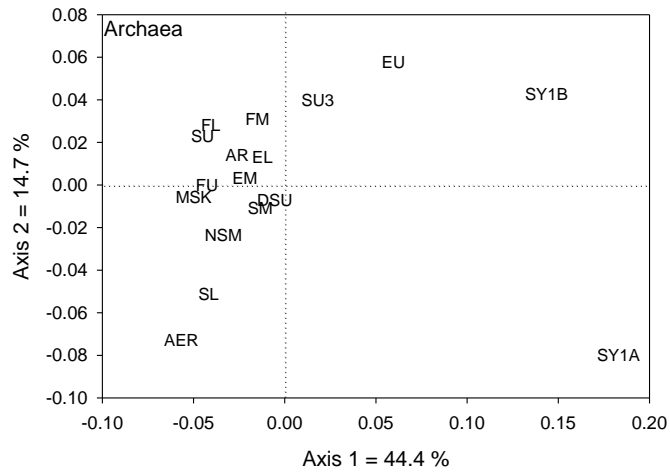


Table S1: Multiplex identifiers used in this study

<i>Sample</i>	<i>MID</i>	<i>Sequence</i>
<i>454</i>		
EU	MID-1	ACGAGTGCGT
EM	MID-2	ACGCTCGACA
EL	MID-3	AGACGCACTC
FU	MID-4	AGCACTGTAG
FM	MID-21	CGTAGACTAG
FL	MID-22	TACGAGTATG
SU	MID-7	CGTGTCTCTA
SM	MID-8	CTCGCGTGTC
SL	MID-10	TCTCTATGCG
AR	MID-11	TGATACGTCT
MSK	MID-13	CATAGTAGTG
SU3	MID-15	ATACGACGTA
NSM	MID-16	TCACGTACTA
SY1A	MID-17	CGTCTAGTAC
SY1B	MID-18	TCTACGTAGC
DSU	MID-19	TGTACTACTC
AER	MID-20	ACGACTACAG
<i>Ion</i>		
<i>Torrent</i>		
EU	MID-1	CTAAGGTAAC
EM	MID-2	TAAGGAGAAC
EL	MID-3	AAGAGGATTC
FU	MID-4	TACCAAGATC
FM	MID-5	CAGAAGGAAC
FL	MID-6	CTGCAAGTTC
SU	MID-7	TTCGTGATTC
SM	MID-8	TTCCGATAAC
SL	MID-9	TGAGCGGAAC
AR	MID-10	CTGACCGAAC
MSK	MID-11	TCCTCGAATC
SU3	MID-12	TAGGTGGTTC
NSM	MID-13	TCTAACGGAC
SY1A	MID-14	TTGGAGTGTC
SY1B	MID-15	TCTAGAGGTC
DSU	MID-16	TCTGGATGAC
AER	MID-17	TACTCACGAT
<i>Positive control</i>	MID-18	TCGTGTCGCA

15 Supplementary Table S2: Number of classified sequences for all the bacterial and *Archaea* phyla (and *Proteobacteria* classes)

16 detected (derived from 454 sequencing for *Bacteria* and from Ion Torrent sequencing for *Archaea*)

Taxa	AER	NSM	DSU	MSK	AR	SU	SM	SL	EU	EM	EL	FU	FM	FL	SU3	SY1A	SY1B
<i>Bacteria</i>																	
Total classified reads	23939	21657	31332	21506	22552	25949	30974	24285	32868	28166	34324	25399	30696	26442	28816	25188	26442
<i>Acidobacteria</i>	1308	429	1219	724	1202	1739	1025	979	607	676	911	818	2004	1303	449	0	5
<i>Actinobacteria</i>	744	336	998	404	556	1905	1103	844	579	316	749	494	903	708	669	124	103
<i>Bacteroidetes</i>	2950	4009	3521	4594	3612	1525	6801	5365	5678	5155	8338	2498	6862	1915	4211	860	1894
<i>BRC1</i>	2	1	3	5	2	1	0	0	0	3	6	0	23	0	1	0	0
<i>Caldiserica</i>	0	0	0	0	0	1	0	0	1	0	0	0	1	1	2	0	7
<i>Chlamydiae</i>	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
<i>Chlorobi</i>	14	5	12	18	19	6	1	1	2	3	11	1	9	2	5	0	0
<i>Chloroflexi</i>	2605	1769	1670	1867	2056	2393	1506	1120	1351	1795	3357	1610	559	718	1101	139	654
<i>Cyanobacteria</i>	883	3395	2883	2230	878	152	544	132	2159	5267	646	6267	111	4443	1327	1	0
<i>Deferribacteres</i>	5	1	2	3	2	0	1	1	1	0	3	1	0	0	5	0	14
<i>Deinococcus-Thermus</i>	12	1	2	3	5	1	1	9	0	2	0	0	0	4	2	18	16
<i>Fibrobacteres</i>	0	0	0	0	0	1	32	12	27	5	3	2	9	3	0	0	0
<i>Firmicutes</i>	420	273	459	435	408	679	1150	3358	5310	387	2030	316	1632	595	391	265	2227
<i>Fusobacteria</i>	7	0	0	1	0	0	0	6	9	7	11	10	31	36	0	0	0
<i>Gemmatimonadetes</i>	193	101	242	118	149	124	197	110	74	190	124	104	372	240	84	0	0
<i>Lentisphaerae</i>	5	8	32	3	1	2	1	0	15	32	9	1	20	12	6	10	5
<i>Nitrospira</i>	106	38	352	78	173	325	99	65	24	54	46	392	365	295	24	1	1
<i>OD1</i>	11	15	62	31	38	16	11	15	9	23	34	9	19	32	50	3	89
<i>OP10</i>	55	8	35	37	45	30	14	16	25	17	18	35	137	29	14	0	4
<i>OP11</i>	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1
<i>Planctomycetes</i>	272	113	340	206	284	125	210	306	142	213	271	344	437	277	130	2	12
<i>Proteobacteria</i>	8862	7664	14315	6408	8797	12085	14372	9303	13252	8714	12984	8729	11607	11827	16822	23303	16106
<i>Epsilonproteobacteria</i>	166	1351	38	29	77	125	191	7	704	156	704	58	619	417	637	569	995
<i>Deltaproteobacteria</i>	2703	1883	3322	2075	2021	1560	3320	4699	5751	1275	3280	1610	1940	1820	1361	2033	6930
<i>Betaproteobacteria</i>	2935	2285	5371	1619	3206	4378	4722	1789	3355	3249	5439	2874	2669	2730	7724	18739	7024
<i>Alphaproteobacteria</i>	1353	747	2374	1429	1743	4198	2149	1452	1717	2272	1304	2250	3887	5062	3717	658	362
<i>Gammaproteobacteria</i>	823	454	1689	605	974	1014	3216	697	746	820	1327	1141	1277	794	2187	1053	213

Taxa	AER	NSM	DSU	MSK	AR	SU	SM	SL	EU	EM	EL	FU	FM	FL	SU3	SY1A	SY1B
<i>Spirochaetes</i>	144	55	57	105	75	7	19	10	19	68	41	28	403	25	89	7	149
<i>SR1</i>	4	19	12	16	17	3	0	5	8	11	11	4	9	28	8	0	0
<i>Synergistetes</i>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	22
<i>Tenericutes</i>	0	2	0	2	0	0	0	0	1	1	0	0	0	0	0	2	29
<i>Thermotogae</i>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	23
<i>TM7</i>	19	24	84	52	47	52	82	53	38	84	57	29	59	129	93	0	33
<i>Verrucomicrobia</i>	789	845	735	610	521	482	777	519	572	1416	639	787	599	553	855	29	22
<i>WS3</i>	11	3	9	4	4	10	1	2	2	2	9	4	11	7	1	0	0
Archaea																	
Total classified reads	718	814	441	546	259	340	781	233	55	160	124	262	176	165	264	334	64
<i>Euryarchaeota</i>	570	506	292	433	186	137	484	87	34	76	66	158	150	111	103	70	63
<i>Crenarchaeota</i>	22	90	30	21	34	98	49	13	4	40	20	39	4	4	12	2	0

17 AR, Athabasca reference site; NSM, Northland sawmill; DSU, downstream Suncor; MSK, Muskeg River; AER, mouth of Ells River;

18 EU, EM, EL: upper, mid and lower Ells River; FU, FM, FL: upper, mid and lower Firebag Creek; SU, SM, SL: upper, mid and lower

19 Steepbank Creek; SY1A, SY1B, Syncrude tailings ponds; SU3: Suncor tailings pond.

20

21

22 Supplementary Table S3: Number of sequences among the 50 most abundant bacterial genera detected in the 454 datasets

Taxa	AER	NSM	DSU	MSK	AR	SU	SM	SL	EU	EM	EL	FU	FM	FL	SU3	SY1A	SY1B
<i>Geobacter</i>	305	202	285	312	227	206	947	2388	2067	176	579	246	402	558	185	367	503
<i>Rhodoferrax</i>	78	48	88	23	43	105	396	121	326	103	383	59	166	241	234	4178	3291
<i>Smithella</i>	420	362	145	380	209	75	22	17	119	25	366	93	71	11	184	44	3247
<i>Flavobacterium</i>	99	69	128	103	43	74	1248	835	487	104	1430	123	103	99	133	51	402
<i>Thiobacillus</i>	142	300	172	59	173	71	190	72	188	157	631	33	0	16	276	2743	28
<i>Clostridium</i>	12	3	32	19	4	199	256	923	2714	13	716	47	153	78	4	0	17
<i>Acidobacteria Gp6</i>	385	121	342	220	390	394	392	285	141	243	283	292	561	432	190	0	0
<i>Verrucomicrobia Subdivision3 genera incertae sedis</i>	377	397	227	259	226	218	278	166	211	568	299	385	278	240	319	0	15
<i>Acidobacteria Gp4</i>	197	37	167	89	207	397	193	250	109	101	196	142	781	297	42	0	5
<i>Paludibacter</i>	60	38	41	61	43	47	435	487	1034	115	215	32	311	81	39	1	30
<i>Nitrospira</i>	106	38	352	78	173	324	99	65	24	54	46	392	365	295	24	1	0
<i>Gemmatimonas</i>	193	101	242	118	149	124	197	110	74	190	124	104	372	240	84	0	0
<i>Leptolinea</i>	233	148	107	169	169	153	121	88	172	152	310	165	78	59	103	1	122
<i>Ottowia</i>	150	73	199	62	157	162	405	109	124	107	149	94	282	83	156	5	2
<i>Cyanobacteria GpIIa</i>	234	146	700	286	216	1	4	2	29	300	37	11	2	37	76	0	0
<i>Cyanobacteria GpI</i>	51	46	76	67	44	26	107	49	815	269	97	105	7	236	29	0	0
<i>Acidobacteria Gp16</i>	148	72	298	102	142	442	132	102	65	43	86	77	90	84	44	0	0
<i>Trichococcus</i>	3	1	3	0	0	8	229	1310	262	0	49	1	0	1	0	0	0
<i>Salinimicrobium</i>	16	5	9	14	5	2	378	179	42	5	477	0	0	0	228	257	246
<i>Arcobacter</i>	0	5	8	6	0	30	171	4	504	120	67	14	496	381	6	0	18
<i>Methylibium</i>	152	111	206	96	197	60	61	68	106	197	119	165	54	41	92	15	16
<i>Ideonella</i>	141	132	166	92	229	25	70	75	144	232	196	79	14	30	53	15	2
<i>Longilinea</i>	161	102	134	136	102	166	65	50	66	94	215	142	44	46	80	0	14
<i>Hyphomicrobium</i>	113	54	170	99	152	187	72	71	76	102	67	99	115	192	38	0	1
<i>Luteolibacter</i>	97	165	157	109	48	27	153	107	131	155	53	53	9	53	182	21	0
<i>Anaerolinea</i>	200	73	107	123	237	140	45	60	66	41	187	61	47	32	45	19	15
<i>Wolinella</i>	59	482	12	9	34	5	10	2	24	4	142	13	29	4	258	127	211
<i>Rhodoplanes</i>	57	5	96	33	45	184	54	44	19	25	22	51	330	263	6	0	0
<i>Opitutus</i>	74	111	94	55	62	66	56	31	65	163	53	106	57	68	162	7	0
<i>Curvibacter</i>	37	39	49	34	23	17	78	54	41	80	57	98	45	68	88	175	157

Taxa	AER	NSM	DSU	MSK	AR	SU	SM	SL	EU	EM	EL	FU	FM	FL	SU3	SY1A	SY1B
<i>Desulforhopalus</i>	27	94	355	37	72	31	65	18	81	30	113	44	2	73	64	0	0
<i>Methylobacter</i>	68	26	31	49	19	320	37	38	70	32	67	249	5	42	41	0	0
<i>Haliea</i>	74	81	124	67	92	11	43	34	46	119	44	109	3	57	106	18	0
<i>Desulfuromonas</i>	35	15	19	25	14	4	130	53	188	27	16	2	2	23	54	34	353
<i>Terrimonas</i>	55	36	106	68	89	28	32	59	50	63	118	51	151	15	28	0	0
<i>Spartobacteria genera incertae sedis</i>	56	20	46	36	45	43	68	102	52	120	89	58	143	42	23	0	0
<i>TM7 genera incertae sedis</i>	19	24	84	52	47	52	82	53	38	84	57	29	59	129	93	0	33
<i>Alistipes</i>	86	94	42	106	81	13	28	5	27	79	52	23	189	17	55	4	28
<i>Acidobacteria Gp7</i>	101	18	75	31	82	131	120	57	67	14	63	12	89	45	16	0	0
<i>Levilinea</i>	106	67	70	86	96	22	28	57	46	39	95	24	37	30	28	0	62
<i>Hydrogenophaga</i>	8	3	8	4	6	0	50	10	44	15	23	7	1	38	395	135	144
<i>Polaromonas</i>	10	3	14	1	9	35	110	10	12	5	42	22	9	23	395	126	18
<i>Treponema</i>	95	35	26	68	39	4	15	8	12	39	23	11	316	16	35	5	87
<i>Pseudomonas</i>	4	3	9	5	14	10	558	7	3	4	19	39	23	20	48	7	58
<i>Steroidobacter</i>	47	26	85	32	66	21	28	18	39	27	21	46	305	25	35	3	0
<i>Bradyrhizobium</i>	25	7	33	22	23	123	27	36	18	8	10	57	265	126	4	0	0
<i>Propionivibrio</i>	36	57	15	19	34	20	28	48	78	57	117	26	112	58	47	0	0
<i>Acidobacteria Gp17</i>	80	26	60	44	63	89	50	42	35	17	38	37	76	53	18	0	0
<i>Ilumatobacter</i>	52	48	104	32	51	61	78	45	32	31	34	35	6	27	76	0	0

23 AR, Athabasca reference site; NSM, Northland sawmill; DSU, downstream Suncor; MSK, Muskeg River; AER, mouth of Ells River;

24 EU, EM, EL: upper, mid and lower Ells River; FU, FM, FL: upper, mid and lower Firebag Creek; SU, SM, SL: upper, mid and lower

25 Steepbank Creek; SY1A, SY1B, Syncrude tailings ponds; SU3: Suncor tailings pond.

26