

Supplementary Information

Table S1. Basic information for sample sites including geography, genetic diversity, and invertebrate taxa used in the analysis.

Sample Site	Latitude	Longitude	Area (ha)	Max. depth (m)	Sample size (stickleback)	Genetic Diversity (n)	Invertebrates
Amor Lake	50°9'27"N	125°34'42"W	362.2	49	17	4.53	Bosmina, Chironomid, Ephemeroptera, Gammarus
Blackwater Lake	50°10'39"N	125°35'20"W	37.5	28	18	4.434	Chironomid, Ephemeroptera, Gammarus
Cedar Lake	50°12'13"N	125°34'3"W	31	20	18	2.767	Bosmina, Chironomid, Cyclopoid, Daphnia, Ephemeroptera, Gammarus, Trichoptera
Little Mud Lake	50°12'24"N	125°33'3"W	4.29	12	19	2.798	Calanoid, Daphnia, Hydracarina, Polyphemus
Muskeg Lake	50°12'5"N	125°34'29"W	12.18	3	20	3.444	Bosmina, Chironomid, Cyclopoid, Daphnia, Ephemeroptera, Gammarus, Polyphemus, Trichoptera
Roberts Lake	50°12'55"N	125°32'29"W	161.3	53	19	4.488	Chironomid, Ephemeroptera, Trichoptera
Blackwater Stream	50°9'51"N	125°35'3w"W	N.A.	N.A.	20	N.A.	Bosmina, Chironomid, Ephemeroptera, Harpactacid, Trichoptera
Roberts Stream	50°13'39"N	125°33'21"W	N.A.	N.A.	17	N.A.	Chironomid, Ephemeroptera, Trichoptera
Campbell River Estuary	50°2'50"N	125°15'37"W	N.A.	N.A.	16	N.A.	Bosmina, Cyclopoid, Gammarus
Sayward Estuary	50°22'46"N	125°56'43"W	N.A.	N.A.	18	N.A.	Chironomid, Gammarus, Harpactacid

Table S2. Stickleback sample size at each site.

Sample Site	n
Amor Lake	17
Blackwater Lake	18
Blackwater Stream	20
Campbell River Estuary	18
Cedar Lake	18
Little Mud Lake	19
Muskeg Lake	20
Roberts Lake	19
Roberts Stream	17
Sayward Estuary	16

Table S3. Invertebrate samples and the number of individuals pooled in each sample for sequencing.

Type	Count	Sample.ID	Type	Count	Sample.ID
Bosmina	10	Ced12mic.plank3	Gammarus	10	CRE12mic.invert1
Bosmina	100	Amo12mic.plank1	Gammarus	3	Say12mic.invert1
Bosmina	100	Mus12mic.plank4	Gammarus	4	Ced12mic.invert2
Bosmina	30	BlkS12mic.plank1	Gammarus	5	Amo12mic.invert2
Bosmina	80	CRE12mic.plank1	Gammarus	6	Blk12mic.invert1
Calanoid	100	LMu12mic.plank3	Gammarus	6	Mus12mic.invert2
Chironomid	1	Blk12mic.invert2	Gammarus	6	Rob12mic.invert2
Chironomid	1	BlkS12mic.invert3	Gammarus	7	Say12mic.inver3
Chironomid	1	Ced12mic.invert4	Harpactacoid	100	Say12mic.plank1
Chironomid	1	Mus12mic.invert6	Harpactacoid	5	BlkS12mic.plank2
Chironomid	1	Say12mic.invert2	Hydracarina	5	LMu12mic.invert1
Chironomid	2	Amo12mic.invert3	Polyphemus	100	LMu12mic.plank2
Chironomid	3	Rob12mic.invert4	Polyphemus	100	Mus12mic.plank3
Chironomid	6	RobS12mic.invert1	trichoptera	1	BlkS12mic.invert5
Cyclopoid	100	Ced12mic.plank1	trichoptera	1	Mus12mic.invert5
Cyclopoid	50	Mus12mic.plank2	trichoptera	1	RobS12mic.invert5
Cyclopoid	60	CRE12mic.plank2	trichoptera	14	Rob12mic.invert3
Daphnia	100	LMu12mic.plank1	trichoptera	2	Ced12mic.invert3
Daphnia	100	Mus12mic.plank1			
Daphnia	60	Ced12mic.plank2			
Ephemeroptera	1	Mus12mic.invert7			
Ephemeroptera	11	Ced12mic.invert1			
Ephemeroptera	2	Blk12mic.invert3			
Ephemeroptera	2	RobS12mic.invert3			
Ephemeroptera	5	Amo12mic.invert1			
Ephemeroptera	5	Mus12mic.invert1			
Ephemeroptera	8	BlkS12mic.invert1			

Table S4. Samples were sequenced in five different sequencing runs, with some overlap between runs. (A) The summary statistics from all five runs, including reads from samples for other projects. (B) The total reads from each sample.

(A)

Run ID: SA13009	Run ID: SA13062
Total number of input sequences: 9848793	Total number of input sequences: 17513529
Barcode not in mapping file: 86169	Barcode not in mapping file: 2853209
Read too short after quality truncation: 406	Read too short after quality truncation: 1682
Count of N characters exceeds limit: 7438	Count of N characters exceeds limit: 35173
Illumina quality digit = 0: 0	Illumina quality digit = 0: 0
Barcode errors exceed max: 633903	Barcode errors exceed max: 0
Median sequence length: 251.00	Median sequence length: 250.00
Total number seqs written: 9120877	Total number seqs written: 14608318
Run ID: SA13063	Run ID: SA13069
Total number of input sequences: 17402611	Total number of input sequences: 17983195
Barcode not in mapping file: 15236199	Barcode not in mapping file: 4267597
Read too short after quality truncation: 245	Read too short after quality truncation: 3038
Count of N characters exceeds limit: 3220	Count of N characters exceeds limit: 12072
Illumina quality digit = 0: 0	Illumina quality digit = 0: 0
Barcode errors exceed max: 0	Barcode errors exceed max: 0
Median sequence length: 250.00	Median sequence length: 250.00
Total number seqs written: 2162947	Total number seqs written: 13673142
Run ID: SA13077	
Total number of input sequences: 9505493	
Barcode not in mapping file: 4610394	
Read too short after quality truncation: 276	
Count of N characters exceeds limit: 7329	
Illumina quality digit = 0: 0	
Barcode errors exceed max: 0	
Median sequence length: 250.00	
Total number seqs written: 4887494	

(B)

Sample ID	Total reads	Sample ID	Total reads	Sample ID	Total reads
AMO12mic.1	71986	Blk12mic.20	25468	BlkS12mic.water2	270055
AMO12mic.10	1283	Blk12mic.3	40288	BlkS12mic.water3	289404
AMO12mic.11	100415	Blk12mic.4	24840	BlkS12mic.water4	408789
AMO12mic.12	76390	Blk12mic.5	37969	BlkS12mic.water5	330065
AMO12mic.13	104306	Blk12mic.6	45093	CRE12mic.1	24381
AMO12mic.14	96031	Blk12mic.7	31505	CRE12mic.10	35890
AMO12mic.15	58976	Blk12mic.8	49672	CRE12mic.11	16720
AMO12mic.16	73858	Blk12mic.9	67490	CRE12mic.12	13124
AMO12mic.17	154200	Blk12mic.invert1	225541	CRE12mic.13	54116
AMO12mic.18	94478	Blk12mic.invert2	249861	CRE12mic.14	18336
AMO12mic.19	112126	Blk12mic.invert3	287243	CRE12mic.15	25915
AMO12mic.2	75600	Blk12mic.plank1	102682	CRE12mic.16	27843
AMO12mic.20	110513	Blk12mic.plank3	118599	CRE12mic.17	13983
AMO12mic.21	70813	Blk12mic.water1	290445	CRE12mic.18	10778
AMO12mic.3	96665	Blk12mic.water2	283248	CRE12mic.19	19409
AMO12mic.4	87838	Blk12mic.water3	303042	CRE12mic.2	36796
AMO12mic.6	106904	BlkS12mic.1	48980	CRE12mic.20	13517
AMO12mic.8	144687	BlkS12mic.10	55061	CRE12mic.3	28530
AMO12mic.9	90302	BlkS12mic.11	20222	CRE12mic.4	26517
Amo12mic.invert1	304098	BlkS12mic.12	35177	CRE12mic.5	11579
Amo12mic.invert2	303546	BlkS12mic.13	18255	CRE12mic.6	25430
Amo12mic.invert3	1003	BlkS12mic.14	13248	CRE12mic.7	17909
Amo12mic.plank1	361996	BlkS12mic.15	40569	CRE12mic.8	16504
Amo12mic.plank2	109662	BlkS12mic.16	49441	CRE12mic.9	16330
Amo12mic.plank3	121002	BlkS12mic.17	33262	CRE12mic.invert1	319944
Amo12mic.water1	362299	BlkS12mic.18	15309	CRE12mic.plank1	250697
Amo12mic.water2	290814	BlkS12mic.19	27787	CRE12mic.plank2	4370
Amo12mic.water3	284217	BlkS12mic.2	34076	CRE12mic.plank3	123199
Amo12mic.water4	313463	BlkS12mic.20	21940	CRE12mic.plank4	195332
Amo12mic.water5	286228	BlkS12mic.3	70232	CRE12mic.water1	321880
Blk12mic.1	26413	BlkS12mic.4	27805	CRE12mic.water2	368881
Blk12mic.10	34769	BlkS12mic.5	25806	CRE12mic.water3	300416
Blk12mic.11	32202	BlkS12mic.6	75316	CRE12mic.water4	279315
Blk12mic.12	23951	BlkS12mic.7	100158	Ced12mic.1	65796
Blk12mic.13	51507	BlkS12mic.8	44330	Ced12mic.10	52166
Blk12mic.14	37596	BlkS12mic.9	32820	Ced12mic.11	57684
Blk12mic.15	31582	BlkS12mic.invert1	250540	Ced12mic.12	59163
Blk12mic.16	22468	BlkS12mic.invert3	102243	Ced12mic.13	51319
Blk12mic.17	52524	BlkS12mic.invert5	195825	Ced12mic.14	83988
Blk12mic.18	16693	BlkS12mic.plank1	278922	Ced12mic.15	51415
Blk12mic.19	38220	BlkS12mic.plank2	271032	Ced12mic.16	90462
Blk12mic.2	51887	BlkS12mic.water1	230662	Ced12mic.17	96594

Sample ID	Total reads	Sample ID	Total reads	Sample ID	Total reads
Ced12mic.18	74959	LMu12mic.invert1	515	ROB12mic.10	84891
Ced12mic.19	95246	LMu12mic.plank1	355722	ROB12mic.11	80912
Ced12mic.2	21448	LMu12mic.plank2	193831	ROB12mic.12	93848
Ced12mic.20	197610	LMu12mic.plank3	176892	ROB12mic.13	107671
Ced12mic.3	48893	LMu12mic.water1	286969	ROB12mic.14	80334
Ced12mic.4	60163	LMu12mic.water2	281676	ROB12mic.15	81843
Ced12mic.5	56569	LMu12mic.water3	329115	ROB12mic.16	111312
Ced12mic.6	51142	LMu12mic.water4	439648	ROB12mic.17	83136
Ced12mic.7	52521	MUS12mic.1	73110	ROB12mic.18	77506
Ced12mic.8	39315	MUS12mic.10	99908	ROB12mic.19	71800
Ced12mic.9	32623	MUS12mic.11	79856	ROB12mic.2	191312
Ced12mic.invert1	383326	MUS12mic.12	104682	ROB12mic.20	62035
Ced12mic.invert2	254861	MUS12mic.13	133226	ROB12mic.3	112661
Ced12mic.invert3	197861	MUS12mic.14	67555	ROB12mic.4	92149
Ced12mic.invert4	203437	MUS12mic.15	63155	ROB12mic.5	102717
Ced12mic.plank1	238290	MUS12mic.16	52305	ROB12mic.6	86869
Ced12mic.plank2	221897	MUS12mic.17	103275	ROB12mic.7	84509
Ced12mic.plank3	357751	MUS12mic.18	34122	ROB12mic.8	127239
Ced12mic.water1	287789	MUS12mic.19	74825	ROB12mic.9	46419
Ced12mic.water2	331645	MUS12mic.2	73250	Rob12mic.invert2	180355
Ced12mic.water3	349088	MUS12mic.20	32340	Rob12mic.invert3	132437
Ced12mic.water4	297589	MUS12mic.3	110344	Rob12mic.invert4	982
LMu12mic.1	32153	MUS12mic.4	88602	Rob12mic.water1	280863
LMu12mic.10	28106	MUS12mic.5	108831	Rob12mic.water2	315432
LMu12mic.11	41971	MUS12mic.6	89357	Rob12mic.water3	337327
LMu12mic.12	54633	MUS12mic.7	71707	Rob12mic.water4	266660
LMu12mic.13	30374	MUS12mic.8	73393	RobS12.mic14	732
LMu12mic.14	20480	MUS12mic.9	72909	RobS12mic.1	25728
LMu12mic.15	27453	Mus12mic.invert1	309847	RobS12mic.10	22596
LMu12mic.16	20813	Mus12mic.invert2	243062	RobS12mic.11	36074
LMu12mic.17	13348	Mus12mic.invert5	204801	RobS12mic.12	130940
LMu12mic.18	27757	Mus12mic.invert6	806	RobS12mic.13	46608
LMu12mic.19	15682	Mus12mic.invert7	256964	RobS12mic.14	30205
LMu12mic.2	42587	Mus12mic.plank1	180527	RobS12mic.15	23216
LMu12mic.20	14192	Mus12mic.plank2	284712	RobS12mic.16	48250
LMu12mic.3	38040	Mus12mic.plank3	205810	RobS12mic.17	56007
LMu12mic.4	70667	Mus12mic.plank4	260717	RobS12mic.18	23794
LMu12mic.5	64262	Mus12mic.water1	288597	RobS12mic.19	44740
LMu12mic.6	38358	Mus12mic.water2	368417	RobS12mic.2	18389
LMu12mic.7	39990	Mus12mic.water3	389002	RobS12mic.20	46348
LMu12mic.8	35320	Mus12mic.water4	309239	RobS12mic.3	15257
LMu12mic.9	47834	ROB12mic.1	124039	RobS12mic.4	16330

Sample ID	Total reads	Sample ID	Total reads
RobS12mic.5	12611	SAY12mic.16	68514
RobS12mic.6	31086	SAY12mic.17	47218
RobS12mic.7	17301	SAY12mic.18	34064
RobS12mic.8	28302	SAY12mic.19	69678
RobS12mic.9	36929	SAY12mic.2	50876
RobS12mic.invert1	180429	SAY12mic.20	36125
RobS12mic.invert3	163663	SAY12mic.3	27803
RobS12mic.invert5	290684	SAY12mic.4	35329
RobS12mic.invert6	63976	SAY12mic.5	89164
RobS12mic.plank1	112660	SAY12mic.6	28144
RobS12mic.plank2	83379	SAY12mic.7	23952
RobS12mic.water1	251889	SAY12mic.8	103068
RobS12mic.water2	340148	SAY12mic.9	34757
RobS12mic.water3	272013	Say12mic.invert1	239211
RobS12mic.water4	289970	Say12mic.invert2	1059
SAY12mic.1	106540	Say12mic.invert3	151235
SAY12mic.10	132762	Say12mic.plank1	1354
SAY12mic.11	15306	Say12mic.water1	481887
SAY12mic.12	39405	Say12mic.water2	483214
SAY12mic.13	68558	Say12mic.water3	297989
SAY12mic.14	68827	Say12mic.water4	398807
SAY12mic.15	53049		

Nine most important OTUs driving variation among habitats

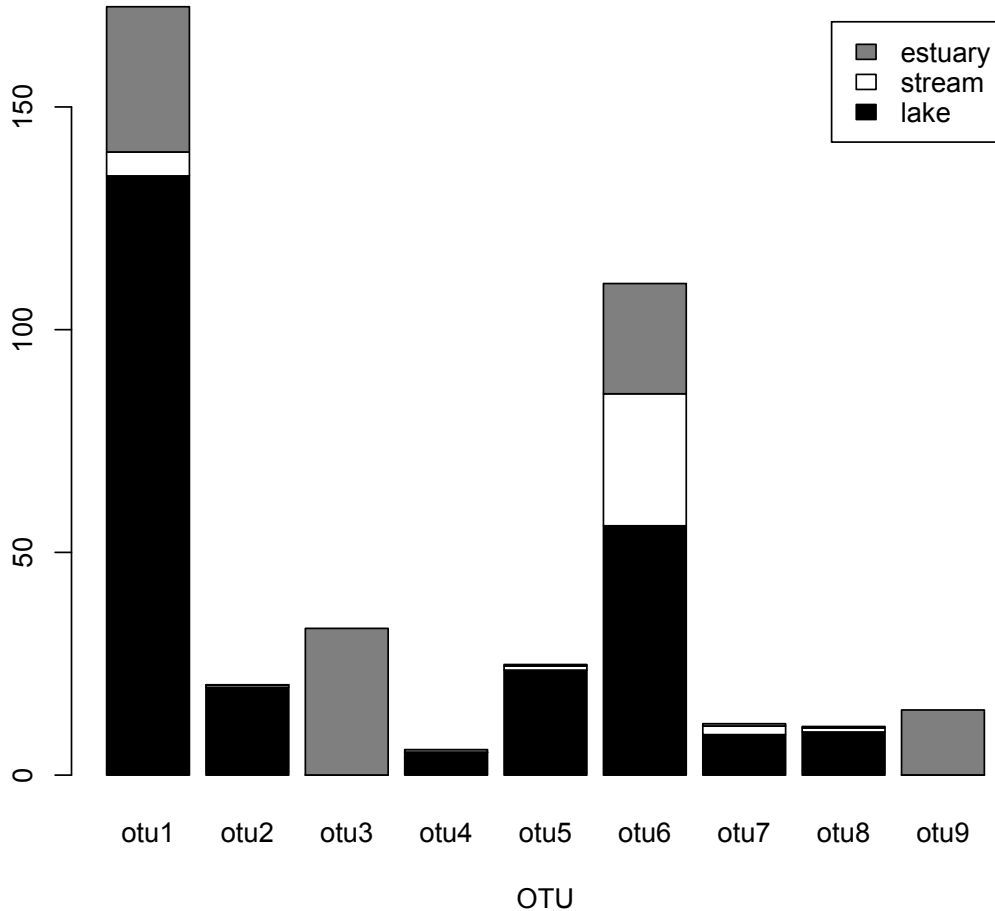


Figure S1. Observed number of counts after rarefaction to 3000 of the nine most important OTUs distinguishing among gut microbiota from different habitat types, calculated using a machine learning algorithm in QIIME. OTU taxonomy summarized in Table S2. Three of the nine most important microbial OTUs distinguishing among habitat types were of the same Cyanobacteria species *Prochlorococcus marinus*, a phytoplankton belonging to possibly the most abundant genus on earth (Flombaum *et al.*, 2013). These OTUs were found mostly in lake fish. For instance, one *P. marinus* OTU

makes up 4.33% of the gut microbiota of lake fish on average, 1% in estuarine fish, and 0.16% in stream fish. On the other hand, some OTUs were found exclusively in estuarine fish, including two Alphaproteobacteria OTUs which make up 0.5-1.0% of marine stickleback gut microbiota.

Table S5. Taxonomy for the nine most influential OTUs contributing to variation among habitats

	Kingdom	Phylum	Class	Order	Family	Genus	Species	OTU #
otu1	Bacteria	Cyanobacter	Synechococ	Synechococ	Synechococ	Prochloroco	marinus	810353
otu2	Bacteria	Cyanobacter	Synechococ	Synechococ	Synechococ	Prochloroco	marinus	4253236
otu3	Bacteria	Proteobacte	Alphaprotec	Rhizobiales	NA	NA	NA	167459
otu4	Bacteria	Cyanobacter	Synechococ	Synechococ	Synechococ	Prochloroco	marinus	557495
otu5	Bacteria	Proteobacte	Deltaproteo	Syntrophob	Syntrophob	NA	NA	701947
otu6	Bacteria	Firmicutes	Bacilli	Bacillales	Alicyclobacil	Alicyclobacil	NA	244966
otu7	Bacteria	Planctomyce	Planctomyce	Pirellulales	Pirellulaceae	NA	NA	2193695
otu8	Bacteria	Planctomyce	Planctomyce	Pirellulales	Pirellulaceae	NA	NA	747323
otu9	Bacteria	Proteobacte	Alphaprotec	Rhodospirilli	Acetobacter	Roseomona	NA	517152

Ten most important OTUs driving variation among lakes

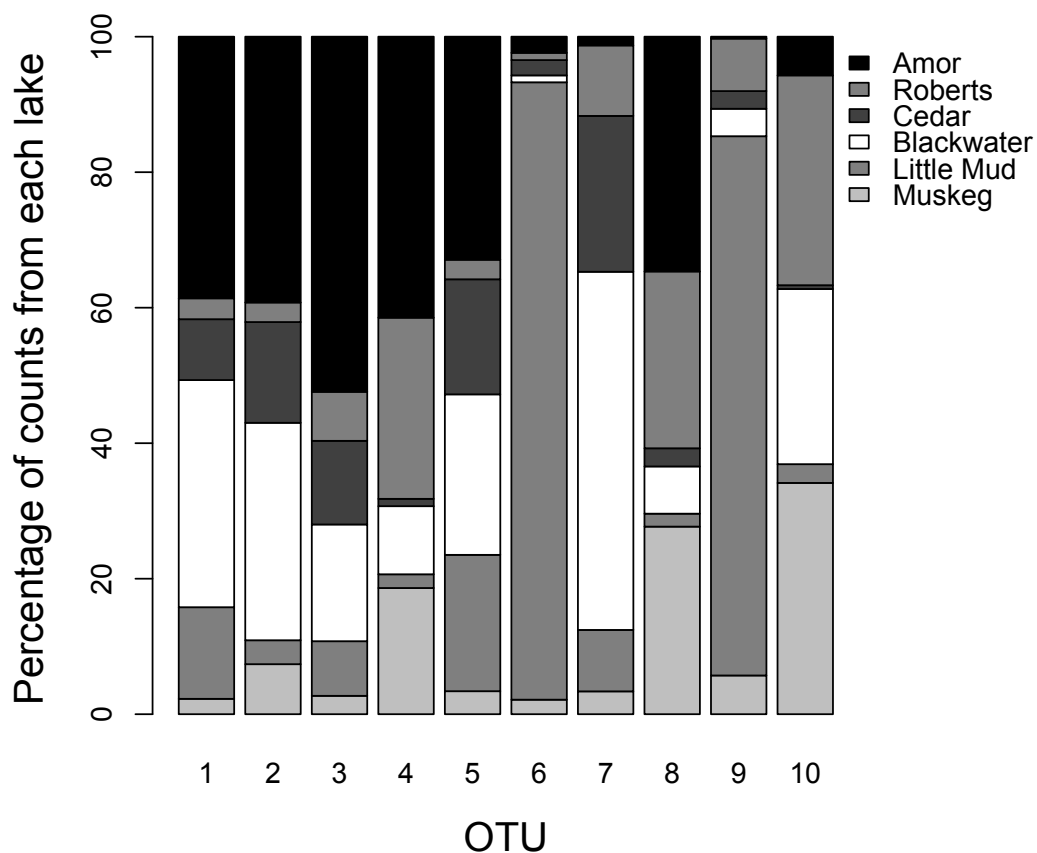


Figure S2. Percentage of total counts after rarefaction to 3000 of the 10 most important OTUs distinguishing among gut microbiota from different lakes, as calculated using QIIME’s supervised_learning.py script. OTU taxonomy summarized in Table S3. One of the top ten OTUs distinguishing among lake populations was the same *P. marinus* that was most influential in distinguishing among habitats (Figure S1). Amor Lake stickleback carried between 3 and 26-fold more *P. marinus* than other lakes. Three of the other top 10 OTUs driving variation among lakes were Proteobacteria of the order Pseudomonadales, and two others were Planctomycetes OTUs in the family Pirellulaceae.

Table S6. Taxonomy for the 10 most influential OTUs contributing to variation among lakes.

Kingdom	Phylum	Class	Order	Family	Genus	Species	OTU #
1 Bacteria	Planctomycete	Planctomycete	Pirellulales	Pirellulaceae	NA	NA	2193695
2 Bacteria	Actinobacteri	Actinobacteri	Actinomycet	Mycobacteri	Mycobacteri	NA	565869
3 Bacteria	Cyanobacter	Synechococ	Synechococ	Synechococ	Prochloroc	marinus	810353
4 Bacteria	Proteobacte	Gammaprot	Pseudomon	Moraxellace	Acinetobact	NA	289967
5 Bacteria	Planctomycete	Planctomycete	Pirellulales	Pirellulaceae	NA	NA	747323
6 Bacteria	Proteobacte	Betaproteot	Burkholderi	Oxalobacter	Janthinobac	lividum	217942
7 Bacteria	Proteobacte	Deltaproteo	Syntrophob	Syntrophob	NA	NA	323943
8 Bacteria	Proteobacte	Gammaprot	Pseudomon	Moraxellace	NA	NA	1141476
9 Bacteria	Proteobacte	Gammaprot	Pseudomon	Pseudomonadaceae	NA	NA	132745
10 Bacteria	Firmicutes	Clostridia	Clostridiales	Peptostrept	NA	NA	344394

Table S7. Results from partial Mantel tests of genetic distance vs microbiota distance (unweighted or weighted UniFrac) among lakes while controlling for geographic distance, with 100,000 permutations.

	unweighted UniFrac	weighted Unifrac
all OTUs	p=0.020, r=0.651	p=0.038, r=0.511
excluding water OTUs	p=0.010, r=0.683	p=0.005, r=0.798
excluding invertebrate OTUs	p=0.003, r=0.778	p=0.006, r=0.769
excluding both water and invertebrate OTUs	p=0.003, r=0.812	p=0.011, r=0.658