

Table S1 - mrkrPOG info2

Table S1 - Taxon-specific marker genes

POG	taxid	description	rank	Quantitative
POG3362	10474	Fuselloviridae	family	Y
POG3538	10477	Lipothrixviridae	family	Y
POG3515	10477	Lipothrixviridae	family	Y
POG3531	10477	Lipothrixviridae	family	Y
POG0337	10656	Tectiviridae	family	Y
POG0573	10662	Myoviridae	family	N
POG0351	10662	Myoviridae	family	N
POG1165	10699	Siphoviridae	family	N
POG0030	10744	Podoviridae	family	N
POG2234	10841	Microviridae	family	Y
POG3628	157897	Rudiviridae	family	N
POG3630	157897	Rudiviridae	family	N
POG3631	157897	Rudiviridae	family	N
POG0337	10657	Tectivirus	genus	Y
POG3354	10475	Fusellovirus	genus	Y
POG3362	10475	Fusellovirus	genus	Y
POG0463	10663	"T4-like viruses"	genus	Y
POG0170	10842	Microvirus	genus	Y
POG0173	10842	Microvirus	genus	Y
POG0174	10842	Microvirus	genus	Y
POG0176	10842	Microvirus	genus	Y
POG0177	10842	Microvirus	genus	Y
POG0178	10842	Microvirus	genus	Y
POG0002	10861	Inovirus	genus	N
POG0346	10875	Plectrovirus	genus	Y
POG0347	10875	Plectrovirus	genus	Y
POG3335	10875	Plectrovirus	genus	N
POG0036	110456	"T7-like viruses"	genus	Y
POG0027	110456	"T7-like viruses"	genus	Y
POG0041	110456	"T7-like viruses"	genus	Y
POG0571	140410	"P2-like viruses"	genus	Y
POG4347	140410	"P2-like viruses"	genus	N
POG1603	186764	"L5-like viruses"	genus	Y
POG1605	186764	"L5-like viruses"	genus	Y
POG1615	186764	"L5-like viruses"	genus	Y
POG2077	186765	"Lambda-like viruses"	genus	N
POG2463	186765	"Lambda-like viruses"	genus	N
POG3564	186765	"Lambda-like viruses"	genus	N
POG3569	186765	"Lambda-like viruses"	genus	N
POG3571	186765	"Lambda-like viruses"	genus	N
POG2793	186777	"Mu-like viruses"	genus	N
POG3328	186777	"Mu-like viruses"	genus	N
POG0302	186794	"P22-like viruses"	genus	Y
POG3488	186794	"P22-like viruses"	genus	N
POG0875	186846	"Phi29-like viruses"	genus	Y
POG0876	186846	"Phi29-like viruses"	genus	Y
POG0878	186846	"Phi29-like viruses"	genus	Y
POG0879	186846	"Phi29-like viruses"	genus	Y
POG0885	186846	"Phi29-like viruses"	genus	Y

Table S1 - mrkrPOG info2

POG2763	187217	"T1-like viruses"	genus	Y
POG2765	187217	"T1-like viruses"	genus	Y
POG2766	187217	"T1-like viruses"	genus	Y
POG2771	187217	"T1-like viruses"	genus	Y
POG2773	187217	"T1-like viruses"	genus	Y
POG2778	187217	"T1-like viruses"	genus	Y
POG2780	187217	"T1-like viruses"	genus	Y
POG2802	187217	"T1-like viruses"	genus	Y
POG3419	308912	"PhiC31-like viruses"	genus	Y
POG3420	308912	"PhiC31-like viruses"	genus	Y
POG3421	308912	"PhiC31-like viruses"	genus	Y
POG3422	308912	"PhiC31-like viruses"	genus	Y
POG3504	341940	Betalipothrixvirus	genus	Y
POG3508	341940	Betalipothrixvirus	genus	Y
POG3508	341940	Betalipothrixvirus	genus	Y
POG2377	371730	"N4-like viruses"	genus	Y
POG2380	371730	"N4-like viruses"	genus	Y
POG2383	371730	"N4-like viruses"	genus	Y
POG2384	371730	"N4-like viruses"	genus	Y
POG2387	371730	"N4-like viruses"	genus	Y
POG2389	371730	"N4-like viruses"	genus	Y
POG2390	371730	"N4-like viruses"	genus	Y
POG3794	477967	"PhiKMV-like viruses"	genus	N
POG1055	542837	"SP6-like viruses"	genus	Y
POG1056	542837	"SP6-like viruses"	genus	Y
POG1057	542837	"SP6-like viruses"	genus	Y
POG1058	542837	"SP6-like viruses"	genus	Y
POG1061	542837	"SP6-like viruses"	genus	Y
POG1064	542837	"SP6-like viruses"	genus	Y
POG1065	542837	"SP6-like viruses"	genus	Y
POG1066	542837	"SP6-like viruses"	genus	Y
POG1191	680114	"I3-like viruses"	genus	Y
POG3252	680115	"PhiKZ-like viruses"	genus	Y
POG3254	680115	"PhiKZ-like viruses"	genus	Y
POG3245	680115	"PhiKZ-like viruses"	genus	Y
POG3732	680287	"AHJD-like viruses"	genus	Y
POG2211	1196844	Hp1likevirus	genus	Y
POG2212	1196844	Hp1likevirus	genus	Y
POG2213	1196844	Hp1likevirus	genus	Y
POG2216	1196844	Hp1likevirus	genus	Y
POG3660	1198138	Bcep781likevirus	genus	Y
POG3661	1198138	Bcep781likevirus	genus	Y
POG3664	1198138	Bcep781likevirus	genus	Y
POG1350	1198140	FelixO1likevirus	genus	Y
POG1375	1198140	FelixO1likevirus	genus	N
POG1425	1198140	FelixO1likevirus	genus	Y
POG1916	1198980	PB1likevirus	genus	Y
POG1919	1198980	PB1likevirus	genus	Y
POG1920	1198980	PB1likevirus	genus	Y
POG1921	1198980	PB1likevirus	genus	Y
POG4377	1198989	PhiCD119likevirus	genus	Y

Table S1 - mrkrPOG info2

POG4378	1198989	PhiCD119likevirus	genus	Y
POG4382	1198989	PhiCD119likevirus	genus	Y
POG4383	1198989	PhiCD119likevirus	genus	Y
POG0252	35237	dsDNA viruses no RNA stage	no rank	N
POG4542	35325	dsRNA nonenveloped viruses	no rank	Y
POG1877	102294	unclassified Caudovirales	no rank	Y
POG1111	196894	unclassified Siphoviridae	no rank	N
POG1151	196894	unclassified Siphoviridae	no rank	N
POG1048	196895	unclassified Podoviridae	no rank	N
POG1355	196896	unclassified Myoviridae	no rank	N
POG0415	329380	unclassified T4-like phages	genus equiv	N
POG0490	329380	unclassified T4-like phages	genus equiv	N
POG3882	349054	unclassified Fusellovirus	genus equiv	Y
POG3884	349054	unclassified Fusellovirus	genus equiv	Y
POG2769	355245	unclassified T1-like viruses	genus equiv	N
POG2785	355245	unclassified T1-like viruses	genus equiv	N
POG4429	356623	unclassified Inovirus	genus equiv	N
POG4430	356623	unclassified Inovirus	genus equiv	N
POG0166	439488	ssRNA viruses	no rank	Y
POG0167	439488	ssRNA viruses	no rank	Y
POG1917	686598	unassigned Myoviridae	no rank	N
POG1924	686598	unassigned Myoviridae	no rank	N
POG0256	686602	unassigned Podoviridae	no rank	N
POG3145	1113536	CBA120-like viruses	no rank	Y
POG3100	1113536	CBA120-like viruses	no rank	Y
POG3116	1113536	CBA120-like viruses	no rank	Y
POG1402	1132574	unclassified Autographivirinae	no rank	Y
POG1406	1132574	unclassified Autographivirinae	no rank	Y
POG0210	28883	Caudovirales	order	N
POG1116	28883	Caudovirales	order	N
POG0744	114416	Lactococcus lactis bacteriophage ul36	species	N
POG0745	114416	Lactococcus lactis bacteriophage ul36	species	N
POG0748	114416	Lactococcus lactis bacteriophage ul36	species	N
POG0750	114416	Lactococcus lactis bacteriophage ul36	species	N
POG0751	114416	Lactococcus lactis bacteriophage ul36	species	N
POG0887	354259	936 group of lactococcal phages	species	Y
POG0893	354259	936 group of lactococcal phages	species	Y
POG0903	354259	936 group of lactococcal phages	species	Y
POG0897	354259	936 group of lactococcal phages	species	Y
POG0820	386445	Bacteriophage 53 sensu lato	species	N
POG0034	542835	Autographivirinae	subfamily	Y
POG0031	542835	Autographivirinae	subfamily	N
POG0644	542836	Nanovirinae	subfamily	Y
POG3382	686565	Gokushovirinae	subfamily	Y
POG0061	857473	Spounavirinae	subfamily	Y
POG0072	857473	Spounavirinae	subfamily	Y
POG0567	857479	Peduovirinae	subfamily	Y
POG0568	857479	Peduovirinae	subfamily	Y
POG0566	857479	Peduovirinae	subfamily	Y
POG0455	1198136	Tevenvirinae	subfamily	Y
POG0172	10239	Vira	undef	N

Table S2 - refG prophages

Table S2 -Prophage regions in reference bacterial genomes that were Identified in th 252 Gut Metagenomic Samples, Their Prevalenc

Bacterial Genome	taxid.NCBIprojectID.locus	Prophage Region (taxid.projid.locus_start_stop)
Bacteroides sp. 1_1_30	457387.68191.NZ_ADCL01000048	457387.68191.NZ_ADCL01000048_186550_252612
Bacteroides sp. 9_1_42FAA	457395.55587.NZ_EQ973128	457395.55587.NZ_EQ973128_174638_241353
Bacteroides sp. 2_2_4	469590.55581.NZ_EQ973359	469590.55581.NZ_EQ973359_342643_366164
Bacteroides ovatus ATCC 8483	411476.54543.NZ_AAXF02000054	411476.54543.NZ_AAXF02000054_17506_58034
Bacteroides stercoris ATCC 43183	449673.54825.NZ_ABFZ02000020	449673.54825.NZ_ABFZ02000020_103161_133419
Bacteroides sp. 20_3	469591.50765.NZ_GG774987	469591.50765.NZ_GG774987_342482_366139
Bacteroides sp. 3_1_33FAA	457391.41705.NZ_GG705227	457391.41705.NZ_GG705227_672728_699329
Bacteroides sp. 9_1_42FAA	457395.55587.NZ_EQ973126	457395.55587.NZ_EQ973126_637669_665814
Alistipes shahii WAL 8301	717959.45913.FP929032	717959.45913.FP929032_169106_212150
Bacteroides stercoris ATCC 43183	449673.54825.NZ_ABFZ02000016	449673.54825.NZ_ABFZ02000016_92495_153157
Bacteroides sp. 1_1_6	469586.55577.NZ_GG695902	469586.55577.NZ_GG695902_90099_130510
Faecalibacterium prausnitzii A2-165	411483.54551.NZ_ACOP02000073	411483.54551.NZ_ACOP02000073_77495_148777
Bacteroides vulgatus ATCC 8482	435590.58253.NC_009614	435590.58253.NC_009614_3600996_3640710
Faecalibacterium prausnitzii A2-165	411483.54551.NZ_ACOP02000009	411483.54551.NZ_ACOP02000009_365_11814
Clostridiales bacterium 1_7_47FAA	457421.55287.NZ_DS990262	457421.55287.NZ_DS990262_142245_182164
Clostridium bolteae ATCC BAA-613	411902.54523.NZ_ABCC02000034	411902.54523.NZ_ABCC02000034_29527_128941
Bacteroides sp. 2_1_16	469587.41707.NZ_GG705209	469587.41707.NZ_GG705209_1211658_1229148
Parabacteroides merdae ATCC 43184	411477.54545.NZ_AAXE02000102	411477.54545.NZ_AAXE02000102_311200_382618
Bacteroides sp. 3_1_23	457390.49963.NZ_GG774949	457390.49963.NZ_GG774949_1661867_1702632
Pseudoflavonifractor capillosus ATCC 29799	411467.54531.NZ_AAXG02000028	411467.54531.NZ_AAXG02000028_63267_221199
Bacteroides sp. 2_1_16	469587.41707.NZ_GG705209	469587.41707.NZ_GG705209_1126479_1197098
Ruminococcus bromii L2-63	657321.39153.FP929051	657321.39153.FP929051_419791_440282
Bifidobacterium longum DJO10A	205913.58833.NC_010816	205913.58833.NC_010816_24635_40803
Bacteroides sp. 3_1_23	457390.49963.NZ_GG774949	457390.49963.NZ_GG774949_1890797_1907195
Bacteroides sp. 3_1_23	457390.49963.NZ_GG774949	457390.49963.NZ_GG774949_1625151_1647307
Eubacterium rectale DSM 17629	657318.39159.FP929042	657318.39159.FP929042_3056410_3114493
Bacteroides sp. 3_1_19	469592.49705.NZ_GG774761	469592.49705.NZ_GG774761_446527_480157
Ruminococcus obeum A2-162	657314.39167.FP929054	657314.39167.FP929054_791554_847791
Butyrivibrio crossotus DSM 2876	511680.55091.NZ_ABWN01000028	511680.55091.NZ_ABWN01000028_5736_32826
Bacteroides fragilis 3_1_12	457424.55575.NZ_EQ973214	457424.55575.NZ_EQ973214_544111_589188
Ruminococcus torques ATCC 27756	411460.54511.NZ_AAVP02000003	411460.54511.NZ_AAVP02000003_28318_88852
Ruminococcus sp. 5_1_39BFAA	457412.55629.NZ_GG696045	457412.55629.NZ_GG696045_305191_333839
Odoribacter splanchnicus DSM 20712	709991.63397.NC_015160	709991.63397.NC_015160_954145_970823
Subdoligranulum variabile DSM 15176	411471.54539.NZ_ACBY02000014	411471.54539.NZ_ACBY02000014_147554_179555
Clostridium cf. saccharolyticum K10	717608.45855.FP929037	717608.45855.FP929037_514091_535651

Table S2 - refG prophages

Bacteroides sp. 3_1_19	469592.49705.NZ_GG774760	469592.49705.NZ_GG774760_646758_661189
Faecalibacterium prausnitzii A2-165	411483.54551.NZ_ACOP02000091	411483.54551.NZ_ACOP02000091_29755_72225
Bacteroides sp. 1_1_6	469586.55577.NZ_GG695899	469586.55577.NZ_GG695899_456277_507973
Faecalibacterium prausnitzii L2-6	718252.45961.FP929045	718252.45961.FP929045_1622989_1691624
Bacteroides ovatus ATCC 8483	411476.54543.NZ_AAXF02000052	411476.54543.NZ_AAXF02000052_228044_266119
Faecalibacterium prausnitzii L2-6	718252.45961.FP929045	718252.45961.FP929045_26945_74593
Faecalibacterium prausnitzii M21/2	411485.54555.NZ_ABED02000029	411485.54555.NZ_ABED02000029_836447_917396
Faecalibacterium prausnitzii M21/2	411485.54555.NZ_ABED02000025	411485.54555.NZ_ABED02000025_159406_265713
Clostridium saccharolyticum WM1	610130.51419.NC_014376	610130.51419.NC_014376_4035542_4048366
Bacteroides sp. 2_2_4	469590.55581.NZ_EQ973359	469590.55581.NZ_EQ973359_483045_538995
Eubacterium rectale ATCC 33656	515619.59169.NC_012781	515619.59169.NC_012781_1139733_1190813
Clostridium leptum DSM 753	428125.54605.NZ_ABCB02000017	428125.54605.NZ_ABCB02000017_74835_193302
Ruminococcus torques L2-14	657313.39169.FP929055	657313.39169.FP929055_82920_135690
Bacteroides sp. 2_1_22	469588.41621.NZ_GG705175	469588.41621.NZ_GG705175_694903_755690
Eubacterium siraeum 70/3	657319.39157.FP929044	657319.39157.FP929044_1599526_1612014
Eubacterium rectale ATCC 33656	515619.59169.NC_012781	515619.59169.NC_012781_1925561_1993269
Bacteroides sp. 20_3	469591.50765.NZ_GG774987	469591.50765.NZ_GG774987_780370_801902
Bacteroides xylanisolvens XB1A	657309.39177.FP929033	657309.39177.FP929033_966470_1018297
Bacteroides fragilis 3_1_12	457424.55575.NZ_EQ973214	457424.55575.NZ_EQ973214_92214_117520
Parabacteroides distasonis ATCC 8503	435591.58301.NC_009615	435591.58301.NC_009615_995121_1045804
Bacteroides sp. 3_1_19	469592.49705.NZ_GG774762	469592.49705.NZ_GG774762_154131_173944
Ruminococcus bromii L2-63	657321.39153.FP929051	657321.39153.FP929051_246701_289340
[Bacteroides] pectinophilus ATCC 43243	483218.54987.NZ_ABVQ01000036	483218.54987.NZ_ABVQ01000036_4827_63278
Roseburia intestinalis XB6B4	718255.45953.FP929050	718255.45953.FP929050_3989067_4039883
Methanobrevibacter smithii ATCC 35061	420247.58827.NC_009515	420247.58827.NC_009515_1697221_1721725
Anaerotruncus colihominis DSM 17241	445972.54807.NZ_ABGD02000003	445972.54807.NZ_ABGD02000003_44922_126081
butyrate-producing bacterium SS3/4	245014.39155.FP929062	245014.39155.FP929062_3015515_3071839
Eubacterium siraeum 70/3	657319.39157.FP929044	657319.39157.FP929044_2141557_2187453
Clostridium cellulovorans 743B	573061.51503.NC_014393	573061.51503.NC_014393_4776423_4835514
Megasphaera sp. UPII 135-E	1000569.71191.NZ_AFUG01000007	1000569.71191.NZ_AFUG01000007_6385_69474
butyrate-producing bacterium SS3/4	245014.39155.FP929062	245014.39155.FP929062_215841_293853
Eubacterium rectale DSM 17629	657318.39159.FP929042	657318.39159.FP929042_1506108_1538396
Gardnerella vaginalis 409-05	553190.43211.NC_013721	553190.43211.NC_013721_977173_999249
butyrate-producing bacterium SSC/2	245018.45957.FP929061	245018.45957.FP929061_50542_88580
Ruminococcus torques ATCC 27756	411460.54511.NZ_AAVP02000011	411460.54511.NZ_AAVP02000011_32895_83826
Clostridiales bacterium 1_7_47FAA	457421.55287.NZ_DS990261	457421.55287.NZ_DS990261_853152_892459
Clostridium bolteae ATCC BAA-613	411902.54523.NZ_ABCC02000033	411902.54523.NZ_ABCC02000033_141849_174342

Table S2 - refG prophages

Ruminococcus torques ATCC 27756	411460.54511.NZ_AAVP02000001	411460.54511.NZ_AAVP02000001_385201_449417
Roseburia intestinalis M50/1	657315.39165.FP929049	657315.39165.FP929049_3659725_3698386
Burkholderiales bacterium 1_1_47	469610.51545.NZ_GL383997	469610.51545.NZ_GL383997_105407_169431
Streptococcus pyogenes MGAS10394	286636.58105.NC_006086	286636.58105.NC_006086_1118971_1173256
Clostridium leptum DSM 753	428125.54605.NZ_ABCB02000018	428125.54605.NZ_ABCB02000018_181402_233637
Bifidobacterium adolescentis L2-32	411481.54549.NZ_AAXD02000028	411481.54549.NZ_AAXD02000028_201419_287098
Ruminococcus sp. SR1/5	657323.39149.FP929053	657323.39149.FP929053_220856_239836
Clostridium cellulolyticum H10	394503.58709.NC_011898	394503.58709.NC_011898_3365305_3418753
Faecalibacterium prausnitzii SL3/3	657322.39151.FP929046	657322.39151.FP929046_2692869_2745505
Dorea formicigenerans ATCC 27755	411461.54513.NZ_AAXA02000009	411461.54513.NZ_AAXA02000009_116431_211230
Burkholderiales bacterium 1_1_47	469610.51545.NZ_GL383997	469610.51545.NZ_GL383997_40963_88345
Clostridium kluyveri DSM 555	431943.58885.NC_009706	431943.58885.NC_009706_2689132_2762463
Clostridium leptum DSM 753	428125.54605.NZ_ABCB02000021	428125.54605.NZ_ABCB02000021_273840_337832
butyrate-producing bacterium SSC/2	245018.45957.FP929061	245018.45957.FP929061_2578900_2630235
Faecalibacterium prausnitzii A2-165	411483.54551.NZ_ACOP02000049	411483.54551.NZ_ACOP02000049_30390_77436
butyrate-producing bacterium SSC/2	245018.45957.FP929061	245018.45957.FP929061_2961127_2984099
Mobiluncus curtisii ATCC 43063	548479.49695.NC_014246	548479.49695.NC_014246_1423112_1479518
Dorea formicigenerans ATCC 27755	411461.54513.NZ_AAXA02000011	411461.54513.NZ_AAXA02000011_114374_187343
Clostridium kluyveri NBRC 12016	583346.59369.NC_011837	583346.59369.NC_011837_2603263_2689130
Faecalibacterium prausnitzii L2-6	718252.45961.FP929045	718252.45961.FP929045_636086_666284
Ruminococcus sp. 5_1_39BFAA	457412.55629.NZ_GG696046	457412.55629.NZ_GG696046_272439_338857
Collinsella aerofaciens ATCC 25986	411903.54525.NZ_AAVN02000001	411903.54525.NZ_AAVN02000001_262289_309092
Shigella flexneri 2a str. 2457T	198215.57991.NC_004741	198215.57991.NC_004741_3385261_3411749
Ruminococcus torques L2-14	657313.39169.FP929055	657313.39169.FP929055_1905717_1947323
Shigella flexneri 2a str. 301	198214.62907.NC_004337	198214.62907.NC_004337_3395178_3421573
Bifidobacterium adolescentis L2-32	411481.54549.NZ_AAXD02000018	411481.54549.NZ_AAXD02000018_1099_13706
Eubacterium siraeum V10Sc8a	717961.45919.FP929059	717961.45919.FP929059_494999_557467
Coprococcus catus GD/7	717962.45861.FP929038	717962.45861.FP929038_2013837_2036392
Bifidobacterium longum DJO10A	205913.58833.NC_010816	205913.58833.NC_010816_1975326_1994586
Streptococcus salivarius CCHSS3	1048332.70481.NC_015760	1048332.70481.NC_015760_22839_44980
Bifidobacterium longum subsp. longum F8	722911.45963.FP929034	722911.45963.FP929034_1919862_1956791
Acidaminococcus intestini RyC-MR95	568816.74445.NC_016077	568816.74445.NC_016077_1256703_1309295
Ruminococcus sp. SR1/5	657323.39149.FP929053	657323.39149.FP929053_1191860_1248849
Bifidobacterium adolescentis L2-32	411481.54549.NZ_AAXD02000052	411481.54549.NZ_AAXD02000052_83678_120800
Streptococcus salivarius CCHSS3	1048332.70481.NC_015760	1048332.70481.NC_015760_1262327_1291039
Victivallis vadensis ATCC BAA-548	340101.54305.NZ_ABDE02000006	340101.54305.NZ_ABDE02000006_80211_128740
Pseudoflavonifractor capillosus ATCC 29799	411467.54531.NZ_AAXG02000002	411467.54531.NZ_AAXG02000002_52750_115090

Table S2 - refG prophages

<i>Shigella flexneri</i> 2a str. 301	198214.62907.NC_004337	198214.62907.NC_004337_4509204_4541262
<i>Acidaminococcus intestini</i> RyC-MR95	568816.74445.NC_016077	568816.74445.NC_016077_1899458_2010316
<i>Bifidobacterium longum</i> subsp. <i>infantis</i> 157F	565040.62693.NC_015052	565040.62693.NC_015052_1206914_1246678
<i>Bacillus cereus</i> B4264	405532.58757.NC_011725	405532.58757.NC_011725_526137_547688
<i>Bifidobacterium longum</i> DJO10A	205913.58833.NC_010816	205913.58833.NC_010816_1321495_1356785
[<i>Bacteroides</i>] <i>pectinophilus</i> ATCC 43243	483218.54987.NZ_ABVQ01000035	483218.54987.NZ_ABVQ01000035_415594_479275
<i>Shigella flexneri</i> 2a str. 301	198214.62907.NC_004337	198214.62907.NC_004337_3605148_3646267
<i>Acidaminococcus intestini</i> RyC-MR95	568816.74445.NC_016077	568816.74445.NC_016077_1331206_1431157
butyrate-producing bacterium SM4/1	245012.45955.FP929060	245012.45955.FP929060_617054_660927
<i>Bifidobacterium longum</i> DJO10A	205913.54103.NZ_AABM02000005	205913.54103.NZ_AABM02000005_6282_33162
<i>Bacteroides fragilis</i> 3_1_12	457424.55575.NZ_EQ973215	457424.55575.NZ_EQ973215_376243_399063
<i>Bifidobacterium bifidum</i> S17	883062.59545.NC_014616	883062.59545.NC_014616_81037_124719
<i>Escherichia coli</i> HS	331112.58393.NC_009800	331112.58393.NC_009800_450152_460516
butyrate-producing bacterium SS3/4	245014.39155.FP929062	245014.39155.FP929062_2620851_2698296
<i>Streptococcus thermophilus</i> LMD-9	322159.58327.NC_008532	322159.58327.NC_008532_469888_497227
<i>Escherichia coli</i> ED1a	585397.59379.NC_011745	585397.59379.NC_011745_2565511_2578087
<i>Escherichia</i> sp. 3_2_53FAA	469598.55601.NZ_DS999459	469598.55601.NZ_DS999459_296146_316146
Clostridiales bacterium 1_7_47FAA	457421.55287.NZ_DS990260	457421.55287.NZ_DS990260_1575043_1608234
<i>Anaerotruncus colihominis</i> DSM 17241	445972.54807.NZ_ABGD02000019	445972.54807.NZ_ABGD02000019_90436_128029
<i>Clostridium bolteae</i> ATCC BAA-613	411902.54523.NZ_ABCC02000048	411902.54523.NZ_ABCC02000048_1037_16159
<i>Subdoligranulum variabile</i> DSM 15176	411471.54539.NZ_ACBY02000054	411471.54539.NZ_ACBY02000054_57416_68041
<i>Bifidobacterium longum</i> NCC2705	206672.57939.NC_004307	206672.57939.NC_004307_1295940_1312867
<i>Escherichia coli</i> ATCC 8739	481805.58783.NC_010468	481805.58783.NC_010468_3719163_3730570
<i>Clostridium bolteae</i> ATCC BAA-613	411902.54523.NZ_ABCC02000001	411902.54523.NZ_ABCC02000001_85613_127632
<i>Shigella sonnei</i> Ss046	300269.58217.NC_007384	300269.58217.NC_007384_2884693_2916444
<i>Escherichia coli</i> HS	331112.58393.NC_009800	331112.58393.NC_009800_3015585_3027405
<i>Escherichia</i> sp. 3_2_53FAA	469598.55601.NZ_DS999462	469598.55601.NZ_DS999462_509895_552140
<i>Bacillus clausii</i> KSM-K16	66692.58237.NC_006582	66692.58237.NC_006582_1439784_1492638
<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhimurium	209261.57973.NC_004631	209261.57973.NC_004631_3501128_3545727
<i>Shigella flexneri</i> 2a str. 2457T	198215.57991.NC_004741	198215.57991.NC_004741_4365829_4405075
<i>Escherichia coli</i> UTI89	364106.58541.NC_007946	364106.58541.NC_007946_897856_958060
<i>Oligotropha carboxidovorans</i> OM5	504832.72795.NC_015684	504832.72795.NC_015684_3295734_3357015
<i>Streptococcus pneumoniae</i> AP200	574093.52453.NC_014494	574093.52453.NC_014494_12370_89253
<i>Escherichia coli</i> 101-1	358709.54363.NZ_AAMK02000001	358709.54363.NZ_AAMK02000001_105168_119197
<i>Streptococcus thermophilus</i> LMG 18311	264199.58219.NC_006448	264199.58219.NC_006448_790359_824358
<i>Bifidobacterium longum</i> subsp. <i>infantis</i> 157F	565040.62693.NC_015052	565040.62693.NC_015052_1006192_1068219
<i>Escherichia coli</i> O103:H2 str. 12009	585395.41013.NC_013353	585395.41013.NC_013353_571967_631372

Table S2 - refG prophages

<i>Escherichia coli</i> 53638	344610.54321.NZ_AAKB02000001	344610.54321.NZ_AAKB02000001_1834929_1875213
<i>Escherichia coli</i> UTI89	364106.58541.NC_007946	364106.58541.NC_007946_1241467_1305787
<i>Escherichia coli</i> IAI39	585057.59381.NC_011750	585057.59381.NC_011750_2735610_2796274
<i>Dehalococcoides ethenogenes</i> 195	243164.57763.NC_002936	243164.57763.NC_002936_961200_1021082
<i>Shigella flexneri</i> 2a str. 301	198214.62907.NC_004337	198214.62907.NC_004337_898668_952154
<i>Escherichia coli</i> 53638	344610.54321.NZ_AAKB02000001	344610.54321.NZ_AAKB02000001_3332186_3386416
<i>Escherichia coli</i> O111:H- str. 11128	585396.41023.NC_013364	585396.41023.NC_013364_2838795_2864257
<i>Escherichia coli</i> B str. REL606	413997.58803.NC_012967	413997.58803.NC_012967_2100409_2143776
<i>Lactobacillus reuteri</i> 100-23	349123.54165.NZ_AAPZ02000001	349123.54165.NZ_AAPZ02000001_397013_461160
<i>Shigella flexneri</i> 2a str. 301	198214.62907.NC_004337	198214.62907.NC_004337_2030765_2095753
<i>Bifidobacterium longum</i> subsp. infantis ATCC 15	391904.58677.NC_011593	391904.58677.NC_011593_1660928_1723489
<i>Escherichia coli</i> IAI39	585057.59381.NC_011750	585057.59381.NC_011750_1882552_1904933
<i>Escherichia coli</i> O26:H11 str. 11368	573235.41021.NC_013361	573235.41021.NC_013361_2516898_2591095
<i>Escherichia coli</i> E24377A	331111.58395.NC_009801	331111.58395.NC_009801_1701981_1714260
<i>Escherichia coli</i> ATCC 8739	481805.58783.NC_010468	481805.58783.NC_010468_3139730_3164528
<i>Ketogulonicigenium vulgare</i> Y25	880591.59581.NC_014625	880591.59581.NC_014625_1944938_2000070
<i>Escherichia coli</i> HS	331112.58393.NC_009800	331112.58393.NC_009800_2798026_2823460
<i>Salmonella enterica</i> subsp. enterica serovar Typh	220341.57793.NC_003198	220341.57793.NC_003198_3494786_3560068
<i>Shigella flexneri</i> 5 str. 8401	373384.58583.NC_008258	373384.58583.NC_008258_1627210_1643338
<i>Desulfovibrio piger</i> ATCC 29098	411464.54519.NZ_ABXU01000033	411464.54519.NZ_ABXU01000033_1_12915
<i>Shigella flexneri</i> 2a str. 2457T	198215.57991.NC_004741	198215.57991.NC_004741_2007446_2067068
<i>Escherichia coli</i> ED1a	585397.59379.NC_011745	585397.59379.NC_011745_2009498_2070958
<i>Veillonella</i> sp. 3_1_44	457416.47845.NZ_GG770201	457416.47845.NZ_GG770201_381043_420806
<i>Anaerotruncus colihominis</i> DSM 17241	445972.54807.NZ_ABGD02000005	445972.54807.NZ_ABGD02000005_245688_293153
<i>Escherichia coli</i> IAI39	585057.59381.NC_011750	585057.59381.NC_011750_1921935_2090060
<i>Shigella flexneri</i> 5 str. 8401	373384.58583.NC_008258	373384.58583.NC_008258_233200_319476
<i>Shigella flexneri</i> 5 str. 8401	373384.58583.NC_008258	373384.58583.NC_008258_1404289_1455338
<i>Anaerotruncus colihominis</i> DSM 17241	445972.54807.NZ_ABGD02000014	445972.54807.NZ_ABGD02000014_1027_105003
<i>Escherichia coli</i> O26:H11 str. 11368	573235.41021.NC_013361	573235.41021.NC_013361_3092601_3155332
<i>Lactococcus lactis</i> subsp. lactis II1403	272623.57671.NC_002662	272623.57671.NC_002662_706695_729241
<i>Lactococcus lactis</i> subsp. cremoris MG1363	416870.58837.NC_009004	416870.58837.NC_009004_695839_705929
<i>Shigella flexneri</i> 2a str. 301	198214.62907.NC_004337	198214.62907.NC_004337_4374732_4399077
<i>Escherichia coli</i> F11	340197.54299.NZ_AAJU02000016	340197.54299.NZ_AAJU02000016_1027_37997
<i>Bacteroides fragilis</i> 3_1_12	457424.55575.NZ_EQ973213	457424.55575.NZ_EQ973213_1647690_1683975
<i>Ruminococcus champanellensis</i> 18P13	213810.39179.FP929052	213810.39179.FP929052_1897458_1922897
<i>Lactococcus lactis</i> subsp. lactis II1403	272623.57671.NC_002662	272623.57671.NC_002662_2066475_2105688
<i>Shigella flexneri</i> 5 str. 8401	373384.58583.NC_008258	373384.58583.NC_008258_4351653_4369898

Table S2 - refG prophages

<i>Escherichia coli</i> CFT073	199310.57915.NC_004431	199310.57915.NC_004431_1394721_1446777
<i>Erysipelotrichaceae bacterium</i> 2_2_44A	457422.73037.NZ_ADCZ01000002	457422.73037.NZ_ADCZ01000002_126177_136812
<i>Escherichia coli</i> SE11	409438.59425.NC_011415	409438.59425.NC_011415_2105912_2167346
<i>Escherichia coli</i> E24377A	331111.58395.NC_009801	331111.58395.NC_009801_2173089_2239760
<i>Escherichia coli</i> ED1a	585397.59379.NC_011745	585397.59379.NC_011745_2381497_2402869
<i>Escherichia coli</i> 53638	344610.54321.NZ_AAKB02000001	344610.54321.NZ_AAKB02000001_361146_386814
<i>Shigella dysenteriae</i> Sd197	300267.13145.NC_007606	300267.13145.NC_007606_2174616_2253854
<i>Desulfovibrio piger</i> ATCC 29098	411464.54519.NZ_ABXU01000029	411464.54519.NZ_ABXU01000029_1_67098
<i>Lactococcus lactis</i> subsp. <i>lactis</i> II1403	272623.57671.NC_002662	272623.57671.NC_002662_615966_632941
<i>Escherichia</i> sp. 3_2_53FAA	469598.55601.NZ_DS999460	469598.55601.NZ_DS999460_731989_788925
<i>Shigella flexneri</i> 2a str. 301	198214.62907.NC_004337	198214.62907.NC_004337_866637_935972
<i>Escherichia coli</i> UTI89	364106.58541.NC_007946	364106.58541.NC_007946_2906465_2962748
<i>Lactococcus lactis</i> subsp. <i>lactis</i> II1403	272623.57671.NC_002662	272623.57671.NC_002662_91052_110385
<i>Anaerotruncus colihominis</i> DSM 17241	445972.54807.NZ_ABGD02000025	445972.54807.NZ_ABGD02000025_58871_159073
<i>Escherichia coli</i> SE11	409438.59425.NC_011415	409438.59425.NC_011415_2261815_2308960
<i>Escherichia coli</i> UTI89	364106.58541.NC_007946	364106.58541.NC_007946_2597497_2638903
<i>Streptococcus</i> sp. C150	435842.62525.NZ_GL698449	435842.62525.NZ_GL698449_1290097_1300970
<i>Escherichia coli</i> E24377A	331111.58395.NC_009801	331111.58395.NC_009801_998185_1019468
<i>Erysipelotrichaceae bacterium</i> 2_2_44A	457422.73037.NZ_ADCZ01000025	457422.73037.NZ_ADCZ01000025_14597_62315
<i>Shigella sonnei</i> Ss046	300269.58217.NC_007384	300269.58217.NC_007384_2927053_2967135
<i>Escherichia coli</i> IAI39	585057.59381.NC_011750	585057.59381.NC_011750_1551760_1590848
<i>Escherichia coli</i> O103:H2 str. 12009	585395.41013.NC_013353	585395.41013.NC_013353_5353463_5416147
<i>Bifidobacterium bifidum</i> PRL2010	702459.59883.NC_014638	702459.59883.NC_014638_1045790_1111236
<i>Escherichia coli</i> ED1a	585397.59379.NC_011745	585397.59379.NC_011745_3016220_3070726
<i>Escherichia coli</i> O111:H- str. 11128	585396.41023.NC_013364	585396.41023.NC_013364_2537511_2667760
<i>Escherichia coli</i> O157:H7 str. TW14588	502346.55087.NZ_ABKY02000001	502346.55087.NZ_ABKY02000001_189500_213178
<i>Escherichia coli</i> IAI39	585057.59381.NC_011750	585057.59381.NC_011750_916451_935209
<i>Clostridium bolteae</i> ATCC BAA-613	411902.54523.NZ_ABCC02000042	411902.54523.NZ_ABCC02000042_4448_65969
<i>Shigella boydii</i> CDC 3083-94	344609.58415.NC_010658	344609.58415.NC_010658_806826_947422
<i>Escherichia coli</i> S88	585035.62979.NC_011742	585035.62979.NC_011742_1181079_1231365
<i>Escherichia coli</i> E110019	340186.54303.NZ_AAJW02000023	340186.54303.NZ_AAJW02000023_17988_70706
<i>Escherichia coli</i> O127:H6 str. E2348/69	574521.59343.NC_011601	574521.59343.NC_011601_710136_770698
<i>Escherichia coli</i> str. K-12 substr. DH10B	316385.58979.NC_010473	316385.58979.NC_010473_1499420_1521094
<i>Escherichia coli</i> O127:H6 str. E2348/69	574521.59343.NC_011601	574521.59343.NC_011601_1268129_1313369
<i>Escherichia coli</i> O26:H11 str. 11368	573235.41021.NC_013361	573235.41021.NC_013361_3035874_3049938
<i>Lactococcus lactis</i> subsp. <i>cremoris</i> SK11	272622.57983.NC_008527	272622.57983.NC_008527_1663287_1692919
<i>Lactococcus lactis</i> subsp. <i>lactis</i> II1403	272623.57671.NC_002662	272623.57671.NC_002662_35516_56120

Table S2 - refG prophages

<i>Escherichia coli</i> ED1a	585397.59379.NC_011745	585397.59379.NC_011745_909097_956824
<i>Escherichia coli</i> O157:H7 str. EC4076	444453.54967.NZ_ABHQ01000015	444453.54967.NZ_ABHQ01000015_11649_108285
<i>Escherichia coli</i> O111:H- str. 11128	585396.41023.NC_013364	585396.41023.NC_013364_1183584_1278700
<i>Escherichia coli</i> O111:H- str. 11128	585396.41023.NC_013364	585396.41023.NC_013364_1730826_1770952
<i>Escherichia coli</i> S88	585035.62979.NC_011742	585035.62979.NC_011742_2798408_2895086
<i>Escherichia coli</i> str. K-12 substr. MG1655	511145.57779.NC_000913	511145.57779.NC_000913_564038_580602
<i>Escherichia coli</i> 53638	344610.54321.NZ_AAKB02000001	344610.54321.NZ_AAKB02000001_3905690_3957429
<i>Escherichia coli</i> str. K-12 substr. DH10B	316385.58979.NC_010473	316385.58979.NC_010473_505693_519934
<i>Escherichia coli</i> S88	585035.62979.NC_011742	585035.62979.NC_011742_558506_610138
<i>Bifidobacterium bifidum</i> NCIMB 41171	398513.55283.NZ_DS990229	398513.55283.NZ_DS990229_965936_1006071
<i>Shigella flexneri</i> 5 str. 8401	373384.58583.NC_008258	373384.58583.NC_008258_1920472_1970425
<i>Escherichia coli</i> IAI39	585057.59381.NC_011750	585057.59381.NC_011750_527470_560126
<i>Yersinia enterocolitica</i> subsp. palearctica 105.5R	994476.63663.NC_015224	994476.63663.NC_015224_2476620_2495330
<i>Escherichia coli</i> E110019	340186.54303.NZ_AAJW02000001	340186.54303.NZ_AAJW02000001_24195_73774
<i>Escherichia coli</i> O127:H6 str. E2348/69	574521.59343.NC_011601	574521.59343.NC_011601_1038706_1086323
<i>Desulfovibrio piger</i> ATCC 29098	411464.54519.NZ_ABXU01000046	411464.54519.NZ_ABXU01000046_1_10055
<i>Escherichia coli</i> CFT073	199310.57915.NC_004431	199310.57915.NC_004431_3021386_3068634
<i>Shigella sonnei</i> Ss046	300269.58217.NC_007384	300269.58217.NC_007384_2536277_2587001
<i>Escherichia coli</i> UMN026	585056.62981.NC_011751	585056.62981.NC_011751_1345931_1452216
<i>Escherichia coli</i> HS	331112.58393.NC_009800	331112.58393.NC_009800_917273_967024
<i>Escherichia coli</i> 101-1	358709.54363.NZ_AAMK02000006	358709.54363.NZ_AAMK02000006_46629_80435
<i>Escherichia coli</i> O26:H11 str. 11368	573235.41021.NC_013361	573235.41021.NC_013361_3614377_3684457
<i>Escherichia coli</i> SMS-3-5	439855.58919.NC_010498	439855.58919.NC_010498_1162007_1215520
<i>Escherichia coli</i> ED1a	585397.59379.NC_011745	585397.59379.NC_011745_1654974_1723756
<i>Escherichia coli</i> HS	331112.58393.NC_009800	331112.58393.NC_009800_2062900_2101925
<i>Lactococcus lactis</i> subsp. <i>lactis</i> II1403	272623.57671.NC_002662	272623.57671.NC_002662_447236_534142
<i>Escherichia coli</i> O157:H7 str. EDL933	155864.57831.NC_002655	155864.57831.NC_002655_2668339_2698542
<i>Escherichia coli</i> O111:H- str. 11128	585396.41023.NC_013364	585396.41023.NC_013364_3317427_3369631
<i>Shigella flexneri</i> 5 str. 8401	373384.58583.NC_008258	373384.58583.NC_008258_932840_942909
<i>Escherichia</i> sp. 3_2_53FAA	469598.55601.NZ_DS999459	469598.55601.NZ_DS999459_1172_42242
<i>Escherichia coli</i> SE11	409438.59425.NC_011415	409438.59425.NC_011415_1475428_1526058
<i>Escherichia coli</i> O157:H7 str. TW14359	544404.59235.NC_013008	544404.59235.NC_013008_3905150_3937999
<i>Escherichia coli</i> O26:H11 str. 11368	573235.41021.NC_013361	573235.41021.NC_013361_2667387_2714788
<i>Escherichia coli</i> O157:H7 str. EC4024	444454.54969.NZ_ABJT01000011	444454.54969.NZ_ABJT01000011_15153_48002
<i>Escherichia coli</i> B str. REL606	413997.58803.NC_012967	413997.58803.NC_012967_777183_795295
<i>Escherichia coli</i> S88	585035.62979.NC_011742	585035.62979.NC_011742_2493355_2534529
<i>Escherichia coli</i> 101-1	358709.54363.NZ_AAMK02000007	358709.54363.NZ_AAMK02000007_82521_97993

Table S2 - refG prophages

<i>Campylobacter hominis</i> ATCC BAA-381	360107.58981.NC_009714	360107.58981.NC_009714_226887_258246
<i>Escherichia coli</i> O111:H- str. 11128	585396.41023.NC_013364	585396.41023.NC_013364_2380989_2440449
<i>Escherichia coli</i> B171	344601.54319.NZ_AAJX02000004	344601.54319.NZ_AAJX02000004_46853_99394
<i>Lactococcus lactis</i> subsp. <i>cremoris</i> MG1363	416870.58837.NC_009004	416870.58837.NC_009004_1226569_1239384
<i>Lactobacillus ruminis</i> ATCC 27782	1069534.73417.NC_015975	1069534.73417.NC_015975_765605_815530
<i>Desulfovibrio piger</i> ATCC 29098	411464.54519.NZ_ABXU01000024	411464.54519.NZ_ABXU01000024_858_36831
<i>Escherichia coli</i> HS	331112.58393.NC_009800	331112.58393.NC_009800_291831_358026
<i>Escherichia coli</i> S88	585035.62979.NC_011742	585035.62979.NC_011742_2169276_2206219
<i>Escherichia coli</i> SMS-3-5	439855.58919.NC_010498	439855.58919.NC_010498_955827_970723
<i>Escherichia coli</i> IAI1	585034.59377.NC_011741	585034.59377.NC_011741_2711042_2765231
<i>Escherichia coli</i> UMN026	585056.62981.NC_011751	585056.62981.NC_011751_1837393_1891630
<i>Escherichia coli</i> UMN026	585056.62981.NC_011751	585056.62981.NC_011751_644943_687075
<i>Escherichia coli</i> 101-1	358709.54363.NZ_AAMK02000005	358709.54363.NZ_AAMK02000005_1841_13342
<i>Lactococcus lactis</i> subsp. <i>lactis</i> II1403	272623.57671.NC_002662	272623.57671.NC_002662_1032978_1090456
<i>Escherichia coli</i> O26:H11 str. 11368	573235.41021.NC_013361	573235.41021.NC_013361_1465781_1521344
<i>Escherichia coli</i> O157:H7 str. TW14588	502346.55087.NZ_ABKY02000005	502346.55087.NZ_ABKY02000005_192313_253457
<i>Escherichia fergusonii</i> ATCC 35469	585054.59375.NC_011740	585054.59375.NC_011740_2752630_2795009
<i>Lactobacillus ruminis</i> SPM0211	1040964.67955.NZ_AFOJ01000007	1040964.67955.NZ_AFOJ01000007_330416_376982
<i>Escherichia coli</i> O26:H11 str. 11368	573235.41021.NC_013361	573235.41021.NC_013361_612093_657903
<i>Escherichia coli</i> O111:H- str. 11128	585396.41023.NC_013364	585396.41023.NC_013364_4952779_5000205
<i>Escherichia coli</i> E110019	340186.54303.NZ_AAJW02000001	340186.54303.NZ_AAJW02000001_305629_316506
<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i> MGH	272620.57619.NC_009648	272620.57619.NC_009648_3798567_3834957
<i>Enterobacter asburiae</i> LF7a	640513.72793.NC_015968	640513.72793.NC_015968_4253610_4268248
<i>Klebsiella variicola</i> At-22	640131.42113.NC_013850	640131.42113.NC_013850_1890403_1946448
<i>Lactobacillus reuteri</i> 100-23	349123.54165.NZ_AAPZ02000001	349123.54165.NZ_AAPZ02000001_118182_176007

Table S2 - refG prophages

Pre and Lytic Activity

start coord	stop coord	NumbSamples	In * Sum Proph	Abund	Numb Samples	PtoH>10
186550	252612	192		19693.91		190
174638	241353	150		18786.52		8
342643	366164	207		16786.32		192
17506	58034	239		11773.17		236
103161	133419	125		10399.42		73
342482	366139	127		10168.89		118
672728	699329	232		9713.15		120
637669	665814	177		8645.84		7
169106	212150	98		6185.76		6
92495	153157	162		5845.98		39
90099	130510	93		5634.61		75
77495	148777	242		4982.09		177
3600996	3640710	186		4603.52		0
365	11814	244		4323.86		174
142245	182164	242		4093.27		239
29527	128941	248		3999.80		229
1211658	1229148	106		3898.77		106
311200	382618	209		3892.53		191
1661867	1702632	142		3415.51		0
63267	221199	242		3151.02		170
1126479	1197098	195		3132.26		167
419791	440282	129		2975.50		0
24635	40803	237		2723.16		229
1890797	1907195	90		2590.20		0
1625151	1647307	100		2562.99		0
3056410	3114493	239		2438.49		6
446527	480157	161		2099.13		5
791554	847791	242		1993.12		17
5736	32826	37		1940.64		0
544111	589188	94		1912.49		78
28318	88852	244		1877.75		228
305191	333839	137		1873.78		59
954145	970823	88		1847.42		13
147554	179555	224		1747.05		41
514091	535651	229		1691.33		127

Table S2 - refG prophages

646758	661189	113	1621.80	0
29755	72225	182	1591.77	31
456277	507973	151	1507.23	43
1622989	1691624	179	1457.61	0
228044	266119	106	1384.40	22
26945	74593	217	1364.13	2
836447	917396	223	1208.99	15
159406	265713	170	1137.39	0
4035542	4048366	194	1133.56	96
483045	538995	119	1132.75	72
1139733	1190813	167	1075.26	0
74835	193302	174	1044.20	174
82920	135690	125	1037.98	11
694903	755690	113	1024.55	88
1599526	1612014	79	989.63	0
1925561	1993269	131	922.43	4
780370	801902	168	901.32	145
966470	1018297	105	873.41	0
92214	117520	16	861.32	15
995121	1045804	136	857.39	5
154131	173944	89	765.59	0
246701	289340	114	708.78	10
4827	63278	78	655.38	39
3989067	4039883	108	634.84	0
1697221	1721725	10	634.19	4
44922	126081	113	632.08	113
3015515	3071839	72	609.35	15
2141557	2187453	63	602.71	0
4776423	4835514	157	551.87	157
6385	69474	143	528.12	0
215841	293853	60	489.25	7
1506108	1538396	86	487.92	0
977173	999249	63	467.80	63
50542	88580	145	448.55	0
32895	83826	74	401.69	50
853152	892459	74	370.76	74
141849	174342	74	365.33	74

Table S2 - refG prophages

385201	449417	111	363.41	100
3659725	3698386	73	327.00	0
105407	169431	69	325.64	25
1118971	1173256	58	306.08	58
181402	233637	61	290.67	42
201419	287098	68	268.56	1
220856	239836	99	266.34	1
3365305	3418753	72	230.50	72
2692869	2745505	69	218.65	0
116431	211230	76	209.46	1
40963	88345	31	200.63	21
2689132	2762463	69	188.61	69
273840	337832	48	187.26	6
2578900	2630235	72	186.97	0
30390	77436	36	175.05	9
2961127	2984099	49	170.35	2
1423112	1479518	61	163.33	61
114374	187343	55	163.28	0
2603263	2689130	62	161.32	62
636086	666284	51	157.80	0
272439	338857	42	152.71	10
262289	309092	30	147.47	8
3385261	3411749	29	136.43	29
1905717	1947323	48	128.85	0
3395178	3421573	26	123.31	26
1099	13706	16	122.15	4
494999	557467	48	121.89	27
2013837	2036392	11	102.43	2
1975326	1994586	27	100.31	14
22839	44980	22	89.92	2
1919862	1956791	11	86.47	6
1256703	1309295	22	77.53	0
1191860	1248849	22	72.00	1
83678	120800	14	71.11	7
1262327	1291039	15	67.53	0
80211	128740	17	61.95	17
52750	115090	11	60.62	0

Table S2 - refG prophages

4509204	4541262	11	59.39	11
1899458	2010316	14	57.90	8
1206914	1246678	21	57.61	6
526137	547688	5	57.24	5
1321495	1356785	9	53.71	5
415594	479275	10	52.65	0
3605148	3646267	11	51.50	11
1331206	1431157	14	50.86	14
617054	660927	15	49.75	2
6282	33162	9	48.51	6
376243	399063	19	47.12	10
81037	124719	6	46.62	6
450152	460516	9	44.92	9
2620851	2698296	11	41.51	0
469888	497227	13	38.47	5
2565511	2578087	7	37.84	7
296146	316146	7	37.76	7
1575043	1608234	11	37.75	4
90436	128029	4	37.09	4
1037	16159	10	36.05	10
57416	68041	5	34.34	0
1295940	1312867	7	33.14	3
3719163	3730570	10	32.86	10
85613	127632	7	32.36	7
2884693	2916444	8	31.82	8
3015585	3027405	7	30.00	7
509895	552140	9	29.08	9
1439784	1492638	2	27.14	2
3501128	3545727	6	26.85	6
4365829	4405075	7	26.62	7
897856	958060	5	26.59	5
3295734	3357015	3	25.77	3
12370	89253	2	25.74	2
105168	119197	4	25.19	4
790359	824358	8	24.76	1
1006192	1068219	8	24.23	4
571967	631372	7	23.94	7

Table S2 - refG prophages

1834929	1875213	7	23.64	7
1241467	1305787	4	23.31	4
2735610	2796274	7	23.07	7
961200	1021082	6	22.70	6
898668	952154	6	21.88	6
3332186	3386416	7	21.86	7
2838795	2864257	6	21.81	6
2100409	2143776	6	20.88	6
397013	461160	3	20.29	3
2030765	2095753	4	20.10	4
1660928	1723489	4	19.94	3
1882552	1904933	7	19.92	7
2516898	2591095	6	19.16	6
1701981	1714260	4	18.67	4
3139730	3164528	6	18.54	6
1944938	2000070	2	18.53	2
2798026	2823460	7	18.50	7
3494786	3560068	4	18.40	4
1627210	1643338	5	17.69	5
1	12915	8	17.65	8
2007446	2067068	4	17.54	4
2009498	2070958	4	17.53	4
381043	420806	2	16.87	2
245688	293153	5	16.83	5
1921935	2090060	5	16.55	5
233200	319476	5	16.49	5
1404289	1455338	5	16.22	5
1027	105003	2	16.00	2
3092601	3155332	4	15.95	4
706695	729241	3	15.65	1
695839	705929	4	15.60	4
4374732	4399077	6	15.47	6
1027	37997	3	15.44	3
1647690	1683975	6	15.34	2
1897458	1922897	3	15.32	0
2066475	2105688	4	15.06	4
4351653	4369898	6	14.97	6

Table S2 - refG prophages

1394721	1446777	6	14.87	6
126177	136812	4	14.83	4
2105912	2167346	4	14.58	4
2173089	2239760	2	14.56	2
2381497	2402869	4	14.53	4
361146	386814	5	14.38	5
2174616	2253854	4	14.34	4
1	67098	5	14.34	5
615966	632941	3	14.30	3
731989	788925	3	14.25	3
866637	935972	5	14.12	3
2906465	2962748	3	13.98	3
91052	110385	3	13.92	3
58871	159073	3	13.79	3
2261815	2308960	2	13.63	2
2597497	2638903	2	13.51	2
1290097	1300970	2	13.26	2
998185	1019468	5	13.25	5
14597	62315	4	13.23	3
2927053	2967135	4	13.09	4
1551760	1590848	4	13.01	4
5353463	5416147	4	12.89	4
1045790	1111236	2	12.86	2
3016220	3070726	2	12.83	2
2537511	2667760	4	12.59	4
189500	213178	5	12.54	5
916451	935209	4	12.38	4
4448	65969	3	12.33	3
806826	947422	4	12.20	4
1181079	1231365	4	12.19	4
17988	70706	3	12.07	3
710136	770698	3	11.99	3
1499420	1521094	4	11.83	4
1268129	1313369	2	11.57	2
3035874	3049938	4	11.41	4
1663287	1692919	3	11.35	3
35516	56120	2	11.26	2

Table S2 - refG prophages

909097	956824	3	11.20	3
11649	108285	3	11.19	3
1183584	1278700	3	11.14	3
1730826	1770952	3	11.00	3
2798408	2895086	3	10.96	3
564038	580602	2	10.90	2
3905690	3957429	3	10.84	3
505693	519934	3	10.73	3
558506	610138	3	10.61	3
965936	1006071	2	10.57	2
1920472	1970425	3	10.44	3
527470	560126	3	10.34	3
2476620	2495330	2	10.23	2
24195	73774	3	10.15	3
1038706	1086323	2	9.96	2
1	10055	5	9.92	5
3021386	3068634	4	9.88	4
2536277	2587001	2	9.60	2
1345931	1452216	4	9.51	3
917273	967024	2	9.41	2
46629	80435	3	9.37	3
3614377	3684457	3	8.95	3
1162007	1215520	4	8.93	2
1654974	1723756	2	8.70	2
2062900	2101925	4	8.63	4
447236	534142	2	8.57	2
2668339	2698542	2	8.50	2
3317427	3369631	2	8.41	2
932840	942909	3	8.25	3
1172	42242	2	8.23	2
1475428	1526058	2	8.10	2
3905150	3937999	3	7.97	3
2667387	2714788	2	7.91	2
15153	48002	3	7.82	3
777183	795295	2	7.76	2
2493355	2534529	2	7.57	2
82521	97993	4	7.38	4

Table S2 - refG prophages

226887	258246	2	7.37	0
2380989	2440449	3	7.07	2
46853	99394	3	6.93	3
1226569	1239384	2	6.78	2
765605	815530	2	6.61	0
858	36831	2	6.58	2
291831	358026	2	6.39	2
2169276	2206219	2	6.39	2
955827	970723	2	6.37	2
2711042	2765231	2	6.30	2
1837393	1891630	2	6.27	2
644943	687075	2	6.11	2
1841	13342	2	6.01	2
1032978	1090456	2	5.78	2
1465781	1521344	2	5.57	2
192313	253457	3	5.31	3
2752630	2795009	3	5.05	1
330416	376982	2	4.98	2
612093	657903	2	4.93	2
4952779	5000205	2	4.85	0
305629	316506	2	4.72	2
3798567	3834957	2	4.69	2
4253610	4268248	2	3.23	2
1890403	1946448	2	3.09	0
118182	176007	2	1.74	0

Table S3 - refG-proph Interact

Table S3 - Reference Bacterial Genomes that were detected in the r
bact.taxic Bacterial Genome

1040964 *Lactobacillus ruminis* SPM0211
 198214 *Shigella flexneri* 2a str. 301
 198215 *Shigella flexneri* 2a str. 2457T
 209261 *Salmonella enterica* subsp. *enterica* serovar Typhi str. Ty2
 220341 *Salmonella enterica* subsp. *enterica* serovar Typhi str. CT18
 272620 *Klebsiella pneumoniae* subsp. *pneumoniae* MGH 78578
 300269 *Shigella sonnei* Ss046
 331112 *Escherichia coli* HS
 331112 *Escherichia coli* HS
 340197 *Escherichia coli* F11
 344610 *Escherichia coli* 53638
 349123 *Lactobacillus reuteri* 100-23
 364106 *Escherichia coli* UTI89
 364106 *Escherichia coli* UTI89
 409438 *Escherichia coli* SE11
 411460 *Ruminococcus torques* ATCC 27756
 411461 *Dorea formicigenerans* ATCC 27755
 411464 *Desulfovibrio piger* ATCC 29098
 411464 *Desulfovibrio piger* ATCC 29098
 411464 *Desulfovibrio piger* ATCC 29098
 411477 *Parabacteroides merdae* ATCC 43184
 411902 *Clostridium bolteae* ATCC BAA-613
 411903 *Collinsella aerofaciens* ATCC 25986
 413997 *Escherichia coli* B str. REL606
 428125 *Clostridium leptum* DSM 753
 428125 *Clostridium leptum* DSM 753
 435591 *Parabacteroides distasonis* ATCC 8503
 445972 *Anaerotruncus colihominis* DSM 17241
 445972 *Anaerotruncus colihominis* DSM 17241
 457412 *Ruminococcus* sp. 5_1_39BFAA
 457421 *Clostridiales bacterium* 1_7_47FAA
 469591 *Bacteroides* sp. 20_3
 469598 *Escherichia* sp. 3_2_53FAA
 469598 *Escherichia* sp. 3_2_53FAA
 469610 *Burkholderiales bacterium* 1_1_47
 469610 *Burkholderiales bacterium* 1_1_47
 568816 *Acidaminococcus intestini* RyC-MR95
 573235 *Escherichia coli* O26:H11 str. 11368
 585034 *Escherichia coli* IAI1
 585035 *Escherichia coli* S88
 585035 *Escherichia coli* S88
 585054 *Escherichia fergusonii* ATCC 35469
 585057 *Escherichia coli* IAI39
 585057 *Escherichia coli* IAI39
 585057 *Escherichia coli* IAI39
 585397 *Escherichia coli* ED1a
 585397 *Escherichia coli* ED1a

Table S3 - refG-proph Interact

metagenomic samples and the phage taxa that were identified as integrated p

Predicted prophage region (taxid.projid.locu_start_st mrkr POG Vir_taxid		
1040964.67955.NZ_AFOJ01000007_330416_376982	POG0252	35237
198214.62907.NC_004337_2030765_2095753	POG3328	186777
198215.57991.NC_004741_2007446_2067068	POG3328	186777
209261.57973.NC_004631_3501128_3545727	POG0571	140410
220341.57793.NC_003198_3494786_3560068	POG0571	140410
272620.57619.NC_009648_3798567_3834957	POG0571	140410
300269.58217.NC_007384_2884693_2916444	POG3328	186777
331112.58393.NC_009800_450152_460516	POG0302	186794
331112.58393.NC_009800_917273_967024	POG0571	140410
340197.54299.NZ_AAJU02000016_1027_37997	POG1924	686598
344610.54321.NZ_AAKB02000001_1834929_1875213	POG0571	140410
349123.54165.NZ_AAPZ02000001_397013_461160	POG0252	35237
364106.58541.NC_007946_1241467_1305787	POG0571	140410
364106.58541.NC_007946_2906465_2962748	POG0302	186794
409438.59425.NC_011415_1475428_1526058	POG0252	35237
411460.54511.NZ_AAVP02000011_32895_83826	POG0252	35237
411461.54513.NZ_AAXA02000011_114374_187343	POG0256	686602
411464.54519.NZ_ABXU01000024_858_36831	POG0573	10662
411464.54519.NZ_ABXU01000033_1_12915	POG0030	10744
411464.54519.NZ_ABXU01000046_1_10055	POG2211	1196844
411477.54545.NZ_AAXE02000102_311200_382618	POG0252	35237
411902.54523.NZ_ABCC02000042_4448_65969	POG0252	35237
411903.54525.NZ_AAVN02000001_262289_309092	POG0252	35237
413997.58803.NC_012967_2100409_2143776	POG0573	10662
428125.54605.NZ_ABCB02000018_181402_233637	POG1111	196894
428125.54605.NZ_ABCB02000021_273840_337832	POG2793	186777
435591.58301.NC_009615_995121_1045804	POG0252	35237
445972.54807.NZ_ABGD02000014_1027_105003	POG0210	28883
445972.54807.NZ_ABGD02000019_90436_128029	POG2793	186777
457412.55629.NZ_GG696046_272439_338857	POG0252	35237
457421.55287.NZ_DS990260_1575043_1608234	POG2211	1196844
469591.50765.NZ_GG774987_780370_801902	POG0252	35237
469598.55601.NZ_DS999460_731989_788925	POG3328	186777
469598.55601.NZ_DS999462_509895_552140	POG0302	186794
469610.51545.NZ_GL383997_105407_169431	POG1917	686598
469610.51545.NZ_GL383997_40963_88345	POG0030	10744
568816.74445.NC_016077_1331206_1431157	POG0252	35237
573235.41021.NC_013361_2667387_2714788	POG0571	140410
585034.59377.NC_011741_2711042_2765231	POG1924	686598
585035.62979.NC_011742_2493355_2534529	POG0302	186794
585035.62979.NC_011742_2798408_2895086	POG2793	186777
585054.59375.NC_011740_2752630_2795009	POG0573	10662
585057.59381.NC_011750_1882552_1904933	POG0571	140410
585057.59381.NC_011750_1921935_2090060	POG0573	10662
585057.59381.NC_011750_2735610_2796274	POG0030	10744
585397.59379.NC_011745_2381497_2402869	POG0573	10662
585397.59379.NC_011745_3016220_3070726	POG0571	140410

Table S3 - refG-proph Interact

rophages

Vir_descrip	Vir_rank
dsDNA viruses, no RNA stage	no rank
Mu-like viruses	genus
Mu-like viruses	genus
P2-like viruses	genus
P2-like viruses	genus
P2-like viruses	genus
Mu-like viruses	genus
P22-like viruses	genus
P2-like viruses	genus
unassigned Myoviridae	no rank
P2-like viruses	genus
dsDNA viruses, no RNA stage	no rank
P2-like viruses	genus
P22-like viruses	genus
dsDNA viruses, no RNA stage	no rank
dsDNA viruses, no RNA stage	no rank
unassigned Podoviridae	no rank
Myoviridae	family
Podoviridae	family
Hp1likevirus	genus
dsDNA viruses, no RNA stage	no rank
dsDNA viruses, no RNA stage	no rank
dsDNA viruses, no RNA stage	no rank
Myoviridae	family
unclassified Siphoviridae	no rank
Mu-like viruses	genus
dsDNA viruses, no RNA stage	no rank
Caudovirales	order
Mu-like viruses	genus
dsDNA viruses, no RNA stage	no rank
Hp1likevirus	genus
dsDNA viruses, no RNA stage	no rank
Mu-like viruses	genus
P22-like viruses	genus
unassigned Myoviridae	no rank
Podoviridae	family
dsDNA viruses, no RNA stage	no rank
P2-like viruses	genus
unassigned Myoviridae	no rank
P22-like viruses	genus
Mu-like viruses	genus
Myoviridae	family
P2-like viruses	genus
Myoviridae	family
Podoviridae	family
Myoviridae	family
P2-like viruses	genus

JCVI_SMPL_GS108a	metagenome	0.1	0.8
JCVI_SMPL_GS108b	metagenome	0.8	3
JCVI_SMPL_GS109	metagenome	0.1	0.8
JCVI_SMPL_GS110a	metagenome	0.1	0.8
JCVI_SMPL_GS110b	metagenome	0.8	3
JCVI_SMPL_GS111	metagenome	0.1	0.8
JCVI_SMPL_GS112a	metagenome	0.1	0.8
JCVI_SMPL_GS112b	metagenome	0.8	3
JCVI_SMPL_GS113	metagenome	0.1	0.8
JCVI_SMPL_GS113	metagenome	0.1	0.8
JCVI_SMPL_GS114	metagenome	0.1	0.8
JCVI_SMPL_GS114	metagenome	0.1	0.8
JCVI_SMPL_GS114	metagenome	0.1	0.8
JCVI_SMPL_GS115	metagenome	0.1	0.8
JCVI_SMPL_GS116	metagenome	0.1	0.8
JCVI_SMPL_GS117a	metagenome	0.1	0.8
JCVI_SMPL_GS117b	metagenome	0.8	3
JCVI_SMPL_GS119	metagenome	0.1	0.8
JCVI_SMPL_GS120	metagenome	0.1	0.8
JCVI_SMPL_GS121	metagenome	0.1	0.8
JCVI_SMPL_GS122a	metagenome	0.1	0.8
JCVI_SMPL_GS122a	metagenome	0.1	0.8
JCVI_SMPL_GS122b	metagenome	0.8	3
JCVI_SMPL_GS123	metagenome	0.1	0.8
JCVI_SMPL_GS148	metagenome	0.1	0.8
JCVI_SMPL_GS148	metagenome	0.1	0.8
JCVI_SMPL_GS149	metagenome	0.1	0.8
JCVI_SMPL_GS149	metagenome	0.1	0.8
cam_pseodu_sample_1	metagenome		

A and analyzed

SAMPLE_DESCRIPTION	SAMPLE_NAME
GS108	GS108
GS112	GS112
GS000a - Sargasso Stations 11 & 13	GS000a
GS000a - Sargasso Stations 11 & 13	GS000a
GS000a - Sargasso Stations 11 & 13	GS000a
GS000a - Sargasso Stations 11 & 13	GS000a
GS000a - Sargasso Stations 11 & 13	GS000a
GS000a - Sargasso Stations 11 & 13	GS000a
GS000a - Sargasso Stations 11 & 13	GS000a
GS000a - Sargasso Stations 11 & 13	GS000a
GS000a - Sargasso Stations 11 & 13	GS000a
GS000a - Sargasso Stations 11 & 13	GS000a
GS000a - Sargasso Stations 11 & 13	GS000a
GS000a - Sargasso Stations 11 & 13	GS000a
GS000b - Sargasso Stations 11 & 13	GS000b
GS000b - Sargasso Stations 11 & 13	GS000b
GS000b - Sargasso Stations 11 & 13	GS000b
GS000b - Sargasso Stations 11 & 13	GS000b
GS000b - Sargasso Stations 11 & 13	GS000b
GS000b - Sargasso Stations 11 & 13	GS000b
GS000b - Sargasso Stations 11 & 13	GS000b
GS000b - Sargasso Stations 11 & 13	GS000b
GS000b - Sargasso Stations 11 & 13	GS000b
GS000b - Sargasso Stations 11 & 13	GS000b
GS000b - Sargasso Stations 11 & 13	GS000b
GS000b - Sargasso Stations 11 & 13	GS000b
GS000b - Sargasso Stations 11 & 13	GS000b
GS000b - Sargasso Stations 11 & 13	GS000b
GS000b - Sargasso Stations 11 & 13	GS000b
GS000c - Sargasso Station 3	GS000c
GS000c - Sargasso Station 3	GS000c
GS000c - Sargasso Station 3	GS000c
GS000d - Sargasso Station 13	GS000d
GS000d - Sargasso Station 13	GS000d
GS000d - Sargasso Station 13	GS000d
GS000d - Sargasso Station 13	GS000d
GS000d - Sargasso Station 13	GS000d
GS000d - Sargasso Station 13	GS000d
GS001a - Hydrostation S	GS001a
GS001a - Hydrostation S	GS001a
GS001b - Hydrostation S	GS001b
GS001c - Hydrostation S	GS001c
GS002 - Gulf of Maine	GS002
GS003 - Browns Bank Gulf of Maine	GS003
GS004 - Outside Halifax Nova Scotia	GS004
GS005 - Bedford Basin Nova Scotia	GS005
GS006 - Bay of Fundy Nova Scotia	GS006
GS007 - Northern Gulf of Maine	GS007
GS008 - Newport Harbor RI	GS008
GS009 - Block Island NY	GS009
GS010 - Cape May NJ	GS010
GS011 - Delaware Bay NJ	GS011
GS012 - Chesapeake Bay MD	GS012

GS012 - Chesapeake Bay MD	GS012
GS013 - Off Nags Head NC	GS013
GS014 - South of Charleston SC	GS014
GS015 - Off Key West FL	GS015
GS016 - Gulf of Mexico	GS016
GS017 - Yucatan Channel	GS017
GS018 - Rosario Bank	GS018
GS019 - Northeast of Colon	GS019
GS020 - Lake Gatun	GS020
GS021 - Gulf of Panama	GS021
GS022 - 250 miles from Panama City	GS022
GS023 - 30 miles from Cocos Island	GS023
GS025 - Dirty Rock Cocos Island	GS025
GS026 - 134 miles NE of Galapagos	GS026
GS026 - 134 miles NE of Galapagos	GS026
GS027 - Devil's Crown Floreana Island	GS027
GS028 - Coastal Floreana	GS028
GS029 - North James Bay Santiago Island	GS029
GS030 - Warm seep Roca Redonda	GS030
GS030 - Warm seep Roca Redonda	GS030
GS031 - Upwelling Fernandina Island	GS031
GS032 - Mangrove on Isabella Island	GS032
GS032 - Mangrove on Isabella Island	GS032
GS033 - Punta Cormorant Floreana Island (Hypersaline Lagoon)	GS033
GS034 - North Seamore Island	GS034
GS035 - Wolf Island	GS035
GS036 - Cabo Marshall Isabella Island	GS036
GS037 - Equatorial Pacific TAO Buoy	GS037
GS038 - Tropical South Pacific	GS038
GS039 - Tropical South Pacific	GS039
GS040 - Tropical South Pacific	GS040
GS041 - Tropical South Pacific	GS041
GS042 - Tropical South Pacific	GS042
GS043 - Tropical South Pacific	GS043
GS044 - 600 miles from F. Polynesia	GS044
GS045 - 400 miles from F. Polynesia	GS045
GS046 - 300 miles from F. Polynesia	GS046
GS047 - 201 miles from F. Polynesia	GS047
GS048a - Moorea Cooks Bay	GS048a
GS049 - Moorea Outside Cooks Bay	GS049
GS050 - Tikehau Lagoon	GS050
GS051 - Rangirora Atoll	GS051
MOVE858 - Chesapeake Bay MD	MOVE858
MOVE858 - Chesapeake Bay MD	MOVE858
MOVE858 - Chesapeake Bay MD	MOVE858
MOVE858 - Chesapeake Bay MD	MOVE858
MOVE858 - Chesapeake Bay MD	MOVE858
MOVE858 - Chesapeake Bay MD	MOVE858
MOVE858 - Chesapeake Bay MD	MOVE858
MOVE858 - Chesapeake Bay MD	MOVE858
GS048b - Inside Cook's Bay Moorea French Polynesia	GS048b

GS108a - Cocos Keeling	Inside Lagoon	GS108a
GS108b - Cocos Keeling	Inside Lagoon	GS108b
GS109 - Indian Ocean		GS109
GS110a - Indian Ocean		GS110a
GS110b - Indian Ocean		GS110b
GS111 - Indian Ocean		GS111
GS112a - Indian Ocean		GS112a
GS112b - Indian Ocean		GS112b
GS113 - Indian Ocean		GS113
GS113 - Indian Ocean		GS113
GS114 - 500 Miles west of the Seychelles in the Indian Ocean		GS114
GS114 - 500 Miles west of the Seychelles in the Indian Ocean		GS114
GS114 - 500 Miles west of the Seychelles in the Indian Ocean		GS114
GS115 - Indian Ocean		GS115
GS116 - Outside Seychelles	Indian Ocean	GS116
GS117a - St. Anne Island	Seychelles	GS117a
GS117b - St. Anne Island	Seychelles	GS117b
GS119 - International Water	Outside of Reunion Island	GS119
GS120 - Madagascar Waters		GS120
GS121 - International water between Madagascar and South Afric		GS121
GS122a - International waters between Madagascar and South Af		GS122a
GS122a - International waters between Madagascar and South Af		GS122a
GS122b - International waters between Madagascar and South Af		GS122b
GS123 - International water between Madagascar and South Afric		GS123
GS148 - East coast Zanzibar (Tanzania)	offshore Paje lagoon	GS148
GS148 - East coast Zanzibar (Tanzania)	offshore Paje lagoon	GS148
GS149 - West coast Zanzibar (Tanzania)	harbour region	GS149
GS149 - West coast Zanzibar (Tanzania)	harbour region	GS149

Pseodo sample for the

MATERIAL_ACC	SITE_NAME	LATITUDE
CAM_BM_GS108	GS108	-12.092617
CAM_BM_GS112	GS112	-8.50525
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS000_S11	Sargasso Station 11	31.175
JCVI_SITE_GS000_S11	Sargasso Station 11	31.175
JCVI_SITE_GS000_S11	Sargasso Station 11	31.175
JCVI_SITE_GS000_S11	Sargasso Station 11	31.175
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS000_S11	Sargasso Station 11	31.175
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS000_S11	Sargasso Station 11	31.175
JCVI_SITE_GS000_S11	Sargasso Station 11	31.175
JCVI_SITE_GS000_S11	Sargasso Station 11	31.175
JCVI_SITE_GS000_S11	Sargasso Station 11	31.175
JCVI_SITE_GS000_S11	Sargasso Station 11	31.175
JCVI_SITE_GS000_S11	Sargasso Station 11	31.175
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS000_S11	Sargasso Station 11	31.175
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS000_S03	Sargasso Stations 3	32.174835
JCVI_SITE_GS000_S03	Sargasso Stations 3	32.174835
JCVI_SITE_GS000_S03	Sargasso Stations 3	32.174835
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS000_S13	Sargasso Station 13	31.175
JCVI_SITE_GS001	Hydrostation S	32.166668
JCVI_SITE_GS001	Hydrostation S	32.166668
JCVI_SITE_GS001	Hydrostation S	32.166668
JCVI_SITE_GS001	Hydrostation S	32.166668
JCVI_SITE_GS002	Gulf of Maine	42.503056
JCVI_SITE_GS003	Browns Bank Gulf of Maine	42.85278
JCVI_SITE_GS004	Outside Halifax Nova Scotia	44.137222
JCVI_SITE_GS005	Bedford Basin Nova Scotia	44.690277
JCVI_SITE_GS006	Bay of Fundy Nova Scotia	45.111668
JCVI_SITE_GS007	Northern Gulf of Maine	43.63222
JCVI_SITE_GS008	Newport Harbor RI	41.485832
JCVI_SITE_GS009	Block Island NY	41.09111
JCVI_SITE_GS010	Cape May NJ	38.94
JCVI_SITE_GS011	Delaware Bay NJ	39.417778
JCVI_SITE_GS012	Chesapeake Bay MD	38.946945

JCVI_SITE_GS012	Chesapeake Bay MD	38.946945
JCVI_SITE_GS013	Off Nags Head NC	36.003887
JCVI_SITE_GS014	South of Charleston SC	32.506943
JCVI_SITE_GS015	Off Key West FL	24.488333
JCVI_SITE_GS016	Gulf of Mexico	24.174723
JCVI_SITE_GS017	Yucatan Channel	20.5225
JCVI_SITE_GS018	Rosario Bank	18.036667
JCVI_SITE_GS019	Northeast of Colon	10.716389
JCVI_SITE_GS020	Lake Gatun	9.164444
JCVI_SITE_GS021	Gulf of Panama	8.129167
JCVI_SITE_GS022	250 miles from Panama City	6.492778
JCVI_SITE_GS023	30 miles from Cocos Island	5.64
JCVI_SITE_GS025	Dirty Rock Cocos Island	5.552778
JCVI_SITE_GS026	134 miles NE of Galapagos	1.2641667
JCVI_SITE_GS026	134 miles NE of Galapagos	1.2641667
JCVI_SITE_GS027	Devil's Crown Floreana Island	-1.2161111
JCVI_SITE_GS028	Coastal Floreana	-1.2169445
JCVI_SITE_GS029	North James Bay Santiago Island	-0.2
JCVI_SITE_GS030	Warm seep Roca Redonda	0.27222222
JCVI_SITE_GS030	Warm seep Roca Redonda	0.27222222
JCVI_SITE_GS031	Upwelling Fernandina Island	-0.3011111
JCVI_SITE_GS032	Mangrove on Isabella Island	-0.5938889
JCVI_SITE_GS032	Mangrove on Isabella Island	-0.5938889
JCVI_SITE_GS033	Punta Cormorant Hypersaline Lagoon Floreana Islan	-1.2283334
JCVI_SITE_GS034	North Seamore Island	-0.38305557
JCVI_SITE_GS035	Wolf Island	1.3891667
JCVI_SITE_GS036	Cabo Marshall Isabella Island	-0.020833334
JCVI_SITE_GS037	Equatorial Pacific TAO Buoy	-1.9738889
JCVI_SITE_GS038	Tropical South Pacific	-2.5819445
JCVI_SITE_GS039	Tropical South Pacific	-3.3433332
JCVI_SITE_GS040	Tropical South Pacific	-4.498889
JCVI_SITE_GS041	Tropical South Pacific	-5.93
JCVI_SITE_GS042	Tropical South Pacific	-7.1075
JCVI_SITE_GS043	Tropical South Pacific	-7.661111
JCVI_SITE_GS044	600 miles from F. Polynesia	-8.415
JCVI_SITE_GS045	400 miles from F. Polynesia	-9.0175
JCVI_SITE_GS046	300 miles from F. Polynesia	-9.571111
JCVI_SITE_GS047	201 miles from F. Polynesia	-10.131389
JCVI_SITE_GS048	Moorea Cooks Bay	-17.475834
JCVI_SITE_GS049	Moorea Outside Cooks Bay	-17.453056
JCVI_SITE_GS050	Tikehau Lagoon	-15.277778
JCVI_SITE_GS051	Rangiroa Atoll	-15.143611
JCVI_SITE_MOVE858	Chesapeake Bay MD	38.133335
JCVI_SITE_MOVE858	Chesapeake Bay MD	38.133335
JCVI_SITE_MOVE858	Chesapeake Bay MD	38.133335
JCVI_SITE_MOVE858	Chesapeake Bay MD	38.133335
JCVI_SITE_MOVE858	Chesapeake Bay MD	38.133335
JCVI_SITE_MOVE858	Chesapeake Bay MD	38.133335
JCVI_SITE_MOVE858	Chesapeake Bay MD	38.133335
JCVI_SITE_MOVE858	Chesapeake Bay MD	38.133335
JCVI_SITE_MOVE858	Chesapeake Bay MD	38.133335
JCVI_SITE_GS048	Moorea Cooks Bay	-17.475834

JCVI_SITE_GS108	Cocos Keeling	Inside Lagoon	-12.0925
JCVI_SITE_GS108	Cocos Keeling	Inside Lagoon	-12.0925
JCVI_SITE_GS109	Indian Ocean		-10.943611
JCVI_SITE_GS110	Indian Ocean		-10.446111
JCVI_SITE_GS110	Indian Ocean		-10.446111
JCVI_SITE_GS111	Indian Ocean		-9.596945
JCVI_SITE_GS112	Indian Ocean		-8.505
JCVI_SITE_GS112	Indian Ocean		-8.505
JCVI_SITE_GS113	Indian Ocean		-7.0075
JCVI_SITE_GS113	Indian Ocean		-7.0075
JCVI_SITE_GS114	500 Miles west of the Seychelles	in the Indian Ocean	-4.990278
JCVI_SITE_GS114	500 Miles west of the Seychelles	in the Indian Ocean	-4.990278
JCVI_SITE_GS114	500 Miles west of the Seychelles	in the Indian Ocean	-4.990278
JCVI_SITE_GS115	Indian Ocean		-4.6625
JCVI_SITE_GS116	Outside Seychelles	Indian Ocean	-4.635
JCVI_SITE_GS117	St. Anne Island	Seychelles	-4.613611
JCVI_SITE_GS117	St. Anne Island	Seychelles	-4.613611
JCVI_SITE_GS119	International Water	Outside of Reunion Island	-23.21611
JCVI_SITE_GS120	Madagascar Waters		-26.035
JCVI_SITE_GS121	International water	between Madagascar and South Africa	-29.348888
JCVI_SITE_GS122	International waters	between Madagascar and South Africa	-30.898333
JCVI_SITE_GS122	International waters	between Madagascar and South Africa	-30.898333
JCVI_SITE_GS122	International waters	between Madagascar and South Africa	-30.898333
JCVI_SITE_GS123	International water	between Madagascar and South Africa	-32.399166
JCVI_SITE_GS148	East coast Zanzibar (Tanzania)	offshore Paje lagoon	-6.3166666
JCVI_SITE_GS148	East coast Zanzibar (Tanzania)	offshore Paje lagoon	-6.3166666
JCVI_SITE_GS149	West coast Zanzibar (Tanzania)	harbour region	-6.116667
JCVI_SITE_GS149	West coast Zanzibar (Tanzania)	harbour region	-6.116667
cam_pseodu_biomaterial_1			

-76.41722	13.2 Estuary	UNITED STATES
-75.39472	2.1 Coastal	UNITED STATES
-79.263885	1 Coastal	UNITED STATES
-83.07	1.7 Coastal	UNITED STATES
-84.344444	2 Coastal Sea	UNITED STATES
-85.41361	2 Open Ocean	MEXICO
-83.78472	1.7 Open Ocean	HONDURAS
-80.25445	1.7 Coastal	PANAMA
-79.83611	2 Fresh Water	PANAMA
-79.69111	1.6 Coastal	PANAMA
-82.90389	2 Open Ocean	PANAMA
-86.56528	2 Open Ocean	COSTA RICA
-87.087776	1.1 Fringing Reef	COSTA RICA
-90.295	2 Open Ocean	ECUADOR
-90.295	2 Open Ocean	ECUADOR
-90.422775	2.2 Coastal	ECUADOR
-90.319725	2 Coastal	ECUADOR
-90.83528	2.1 Coastal	ECUADOR
-91.63333	19 Warm Seep	ECUADOR
-91.63333	19 Warm Seep	ECUADOR
-91.651665	12 Coastal upwelling	ECUADOR
-91.06944	0.1 Mangrove	ECUADOR
-91.06944	0.1 Mangrove	ECUADOR
-90.42917	0.2 Hypersaline	ECUADOR
-90.279724	2.1 Coastal	ECUADOR
-91.81695	1.7 Coastal	ECUADOR
-91.19778	2.1 Coastal	ECUADOR
-95.014725	1.8 Open Ocean	INTERNATIONAL
-97.85139	1.8 Open Ocean	INTERNATIONAL
-101.373886	2 Open Ocean	INTERNATIONAL
-105.07	2.2 Open Ocean	INTERNATIONAL
-108.68694	2 Open Ocean	INTERNATIONAL
-116.11916	1.7 Open Ocean	INTERNATIONAL
-120.40222	1.9 Open Ocean	INTERNATIONAL
-124.23972	2 Open Ocean	INTERNATIONAL
-127.76722	1.7 Open Ocean	INTERNATIONAL
-131.49167	1.9 Open Ocean	INTERNATIONAL
-135.44945	30 Open Ocean	FRENCH POLYNESIA
-149.81223	1.4 Coral Reef	FRENCH POLYNESIA
-149.79889	1.4 Coastal	FRENCH POLYNESIA
-148.22444	1.2 Coral Atoll	FRENCH POLYNESIA
-147.435	1 Coral Reef Atoll	FRENCH POLYNESIA
-76.38333	2.07 Estuary	UNITED STATES
-76.38333	2.07 Estuary	UNITED STATES
-76.38333	2.07 Estuary	UNITED STATES
-76.38333	2.07 Estuary	UNITED STATES
-76.38333	2.07 Estuary	UNITED STATES
-76.38333	2.07 Estuary	UNITED STATES
-76.38333	2.07 Estuary	UNITED STATES
-76.38333	2.07 Estuary	UNITED STATES
-149.81223	1.4 Coral Reef	FRENCH POLYNESIA

96.88167	1.8 Lagoon Reef	AUSTRALIA
96.88167	1.8 Lagoon Reef	AUSTRALIA
92.05889	1.5 Open Ocean	INTERNATIONAL
88.30278	1.5 Open Ocean	INTERNATIONAL
88.30278	1.5 Open Ocean	INTERNATIONAL
84.1975	1.8 Open Ocean	INTERNATIONAL
80.37556	1.8 Open Ocean	INTERNATIONAL
80.37556	1.8 Open Ocean	INTERNATIONAL
76.33139	1.8 Open Ocean	INTERNATIONAL
76.33139	1.8 Open Ocean	INTERNATIONAL
64.97667	1.5 Open Ocean	INTERNATIONAL
64.97667	1.5 Open Ocean	INTERNATIONAL
64.97667	1.5 Open Ocean	INTERNATIONAL
60.523056	1.5 Open Ocean	INTERNATIONAL
56.836113	1.5 Open Ocean	SEYCHELLES
55.50861	1.8 Coastal sample	SEYCHELLES
55.50861	1.8 Coastal sample	SEYCHELLES
52.30611	2 Open Ocean	INTERNATIONAL
50.123055	2.8 Open Ocean	MADAGASCAR
43.215557	1.5 Open Ocean	INTERNATIONAL
40.420277	1.9 Open Ocean	INTERNATIONAL
40.420277	1.9 Open Ocean	INTERNATIONAL
40.420277	1.9 Open Ocean	INTERNATIONAL
36.591946	2.2 Open Ocean	INTERNATIONAL
39.009167	0.3 Fringing Reef	TANZANIA UNITED REPUBLIC OF
39.009167	0.3 Fringing Reef	TANZANIA UNITED REPUBLIC OF
39.116665	1.5 Harbor	TANZANIA UNITED REPUBLIC OF
39.116665	1.5 Harbor	TANZANIA UNITED REPUBLIC OF

REGION	HABITAT_NAME	LIBRARY_ACC
Coccos Keeling	Inside Lago	marine habitat CAM_LIB_GOS108XLRVAL-4F-1-400
Indian Ocean		marine habitat CAM_LIB_GOS112XLRVAL-4F-1-400
Sargasso Sea		saline water JCVI_LIB_GS-00a-01-01-9P0KB-T12947
Sargasso Sea		saline water JCVI_LIB_GS-00a-01-01-3P5KB-T12946
Sargasso Sea		saline water JCVI_LIB_GS-00a-01-01-10P0KB-T12948
Sargasso Sea		saline water JCVI_LIB_GS-00a-01-01-2P5KB-T13532
Sargasso Sea		saline water JCVI_LIB_GS-00a-01-01-2P0KB-T13512
Sargasso Sea		saline water JCVI_LIB_GS-00a-01-01-3P0KB-T12945
Sargasso Sea		saline water JCVI_LIB_GS-00a-01-01-3P0KB-T12945
Sargasso Sea		saline water JCVI_LIB_GS-00a-01-01-2P0KB-T13512
Sargasso Sea		saline water JCVI_LIB_GS-00a-01-01-2P5KB-T13532
Sargasso Sea		saline water JCVI_LIB_GS-00a-01-01-10P0KB-T12948
Sargasso Sea		saline water JCVI_LIB_GS-00a-01-01-3P5KB-T12946
Sargasso Sea		saline water JCVI_LIB_GS-00a-01-01-9P0KB-T12947
Sargasso Sea		saline water JCVI_LIB_GS-00b-01-01-9P0KB-T12888
Sargasso Sea		saline water JCVI_LIB_GS-00b-01-01-10P0KB-T12889
Sargasso Sea		saline water JCVI_LIB_GS-00b-01-01-1P6-2P0KB-T12889
Sargasso Sea		saline water JCVI_LIB_GS-00b-01-01-2P0-2P5KB-T13513
Sargasso Sea		saline water JCVI_LIB_GS-00b-01-01-3P0-4P0KB-T12889
Sargasso Sea		saline water JCVI_LIB_GS-00b-01-01-9P0KB-T12888
Sargasso Sea		saline water JCVI_LIB_GS-00b-01-01-10P0KB-T13513
Sargasso Sea		saline water JCVI_LIB_GS-00b-01-01-10P0KB-T12889
Sargasso Sea		saline water JCVI_LIB_GS-00b-01-01-1P6-2P0KB-T12889
Sargasso Sea		saline water JCVI_LIB_GS-00b-01-01-10P0KB-T13513
Sargasso Sea		saline water JCVI_LIB_GS-00b-01-01-3P0-4P0KB-T12889
Sargasso Sea		saline water JCVI_LIB_GS-00b-01-01-2P0-2P5KB-T13513
Sargasso Sea		saline water JCVI_LIB_GS-00c-01-01-9P0KB-T13751
Sargasso Sea		saline water JCVI_LIB_GS-00c-01-01-3P5KB-T13750
Sargasso Sea		saline water JCVI_LIB_GS-00c-01-01-10P0KB-T13752
Sargasso Sea		saline water JCVI_LIB_GS-00d-01-01-9P0KB-T13756
Sargasso Sea		saline water JCVI_LIB_GS-00d-01-01-2P0KB-T13753
Sargasso Sea		saline water JCVI_LIB_GS-00d-01-01-2P5KB-T13754
Sargasso Sea		saline water JCVI_LIB_GS-00d-01-01-10P0KB-T13757
Sargasso Sea		saline water JCVI_LIB_GS-00d-01-01-3P5KB-T13755
Sargasso Sea		saline water JCVI_LIB_GS-01a-01-01-6P0KB-T13985
Sargasso Sea		saline water JCVI_LIB_GS-01a-01-01-2P0KB-T13984
Sargasso Sea		saline water JCVI_LIB_GS-01b-01-01-2P0-4P0KB-T13984
Sargasso Sea		saline water JCVI_LIB_GS-01c-01-01-3P0KB-T13940
North American East Coast		saline water JCVI_LIB_GS-02-01-01-1P6KB
North American East Coast		saline water JCVI_LIB_GS-03-01-01-2P2KB
North American East Coast		saline water JCVI_LIB_GS-04-01-01-1P4-1P8KB
North American East Coast		saline water JCVI_LIB_GS-05-01-01-1P6KB
North American East Coast		saline water JCVI_LIB_GS-06-01-01-1P8KB
North American East Coast		saline water JCVI_LIB_GS-07-01-01-1P6KB
North American East Coast		saline water JCVI_LIB_GS-08-01-01-1P6KB
North American East Coast		saline water JCVI_LIB_GS-09-01-01-1P2-1P4KB
North American East Coast		saline water JCVI_LIB_GS-10-01-01-1P5KB
North American East Coast		saline water JCVI_LIB_GS-11-01-01-1P3-1P8KB
North American East Coast		saline water JCVI_LIB_GS-12-01-01-1P3-1P6KB

North American East Coast	saline water	JCVI_LIB_GS-12-02-01-1P6-2P0KB
North American East Coast	saline water	JCVI_LIB_GS-13-02-01-1P6-2P0KB
North American East Coast	saline water	JCVI_LIB_GS-14-01-01-1P6-2P0KB
Caribbean Sea	saline water	JCVI_LIB_GS-15-01-01-1P3-1P6KB
Caribbean Sea	saline water	JCVI_LIB_GS-16-01-01-1P3-1P6KB
Caribbean Sea	saline water	JCVI_LIB_GS-17-01-01-1P8-2P2KB
Caribbean Sea	saline water	JCVI_LIB_GS-18-01-01-1P6-2P0KB
Caribbean Sea	saline water	JCVI_LIB_GS-19-01-01-1P3-1P6KB
Panama Canal	freshwater habitat	JCVI_LIB_GS-20-01-01-1P3-1P8KB
Eastern Tropical Pacific	saline water	JCVI_LIB_GS-21-01-01-1P3-1P8KB
Eastern Tropical Pacific	saline water	JCVI_LIB_GS-22-01-01-1P3-1P6KB
Eastern Tropical Pacific	saline water	JCVI_LIB_GS-23-01-01-1P8-2P2KB
Eastern Tropical Pacific	saline water	JCVI_LIB_GS-25-01-01-1P8-2P2KB
Galapagos Islands	saline water	JCVI_LIB_GS-26-01-01-1P3-1P6KB
Galapagos Islands	saline water	JCVI_LIB_GS-26-02-01-1P0-1P3KB
Galapagos Islands	saline water	JCVI_LIB_GS-27-01-01-1P8-2P0KB
Galapagos Islands	saline water	JCVI_LIB_GS-28-01-01-1P6-2P0KB
Galapagos Islands	saline water	JCVI_LIB_GS-29-01-01-1P0-1P3KB
Galapagos Islands	saline water	JCVI_LIB_GS-30-01-01-1P3-1P6KB
Galapagos Islands	saline water	JCVI_LIB_GS-30-02-01-1P0-1P3KB
Galapagos Islands	saline water	JCVI_LIB_GS-31-01-01-1P3-1P8KB
Galapagos Islands	saline water	JCVI_LIB_GS-32-01-01-1P3-1P6KB
Galapagos Islands	saline water	JCVI_LIB_GS-32-02-01-1P0-1P3KB
Galapagos Islands	saline water	JCVI_LIB_GS-33-01-01-1P3-1P8KB
Galapagos Islands	saline water	JCVI_LIB_GS-34-01-01-1P5-1P8KB
Galapagos Islands	saline water	JCVI_LIB_GS-35-01-01-1P5KB
Galapagos Islands	saline water	JCVI_LIB_GS-36-01-01-2P2KB
Eastern Tropical Pacific	saline water	JCVI_LIB_GS-37-01-01-1P5KB
Tropical South Pacific	saline water	JCVI_LIB_GS-38-01-01-1P5KB
Tropical South Pacific	saline water	JCVI_LIB_GS-39-01-01-2P0KB
Tropical South Pacific	saline water	JCVI_LIB_GS-40-01-01-1P6-2P0KB
Tropical South Pacific	saline water	JCVI_LIB_GS-41-01-01-1P6-2P0KB
Tropical South Pacific	saline water	JCVI_LIB_GS-42-01-01-1P6-2P0KB
Tropical South Pacific	saline water	JCVI_LIB_GS-43-01-01-1P6-2P0KB
Tropical South Pacific	saline water	JCVI_LIB_GS-44-01-01-1P6-2P0KB
Tropical South Pacific	saline water	JCVI_LIB_GS-45-01-01-1P6-2P0KB
Tropical South Pacific	saline water	JCVI_LIB_GS-46-01-01-1P0-1P3KB
Tropical South Pacific	saline water	JCVI_LIB_GS-47-01-01-1P6-2P0KB
Polynesia Archipelagos	saline water	JCVI_LIB_GS-48-01-01-1P6-2P0KB
Polynesia Archipelagos	saline water	JCVI_LIB_GS-49-01-01-1P6-2P0KB
Polynesia Archipelagos	saline water	JCVI_LIB_GS-50-01-01-1P3-1P6KB
Polynesia Archipelagos	saline water	JCVI_LIB_GS-51-01-01-1P6-2P0KB
North American East Coast	saline water	JCVI_LIB_MOVE858-B
North American East Coast	saline water	JCVI_LIB_MOVE858-A
North American East Coast	saline water	JCVI_LIB_MOVE858-H
North American East Coast	saline water	JCVI_LIB_MOVE858-G
North American East Coast	saline water	JCVI_LIB_MOVE858-F
North American East Coast	saline water	JCVI_LIB_MOVE858-E
North American East Coast	saline water	JCVI_LIB_MOVE858-D
North American East Coast	saline water	JCVI_LIB_MOVE858-C
Polynesia Archipelagos	saline water	JCVI_LIB_GSIOSM048-G-01-8-10KB

Indian Ocean	saline water	JCVI_LIB_GS108-G-02-4-6KB
Indian Ocean	saline water	JCVI_LIB_GSIOSM108-G-01-3-4KB
Indian Ocean	saline water	JCVI_LIB_GS109-G-01-3-4KB
Indian Ocean	saline water	JCVI_LIB_GS110-G-02-4-6KB
Indian Ocean	saline water	JCVI_LIB_GSIOSM110-G-01-3-4KB
Indian Ocean	saline water	JCVI_LIB_GS111-G-02-4-6KB
Indian Ocean	saline water	JCVI_LIB_GS112-G-02-4-6KB
Indian Ocean	saline water	JCVI_LIB_GSIOSM112-G-01-3-4KB
Indian Ocean	saline water	JCVI_LIB_GS113-G-01-4-6KB
Indian Ocean	saline water	JCVI_LIB_GS113-G-02-6-8KB
Indian Ocean	saline water	JCVI_LIB_GS114-G-02-4-6KB
Indian Ocean	saline water	JCVI_LIB_GS114-G-04-1P6-2KB
Indian Ocean	saline water	JCVI_LIB_GS114-G-03-2-3KB
Indian Ocean	saline water	JCVI_LIB_GS115-G-02-4-6KB
Indian Ocean	saline water	JCVI_LIB_GS116-G-02-4-6KB
Indian Ocean	saline water	JCVI_LIB_GS117-G-02-3-5KB
Indian Ocean	saline water	JCVI_LIB_GSIOSM117-G-01-3-4KB
Indian Ocean	saline water	JCVI_LIB_GS119-G-01-3-4KB
Indian Ocean	saline water	JCVI_LIB_GS120-G-02-4-6KB
Indian Ocean	saline water	JCVI_LIB_GS121-G-01-4-6KB
Indian Ocean	saline water	JCVI_LIB_GS122-G-02-4-6KB
Indian Ocean	saline water	JCVI_LIB_GS122-G-4-6kb
Indian Ocean	saline water	JCVI_LIB_GSIOSM122-G-01-3-4KB
Indian Ocean	saline water	JCVI_LIB_GS123-G-01-4-6KB
Indian Ocean	saline water	JCVI_LIB_GS148-G-01-4-6kb
Indian Ocean	saline water	JCVI_LIB_GS148-G-02-6-8kb
Indian Ocean	saline water	JCVI_LIB_GS149-G-01-4-6kb
Indian Ocean	saline water	JCVI_LIB_GS149-G-02-6-8kb

SEQUENCING_METHOD	DNA_TYPE	NUM_OF_READS	MATERIAL_ID
pyrosequencing (454)	gDNA	1382197	1401
pyrosequencing (454)	gDNA	1156475	1402
dideoxysequencing (Sanger gDNA		237804	86
dideoxysequencing (Sanger gDNA		167	84
dideoxysequencing (Sanger gDNA		231338	84
dideoxysequencing (Sanger gDNA		20768	84
dideoxysequencing (Sanger gDNA		154297	84
dideoxysequencing (Sanger gDNA		177	86
dideoxysequencing (Sanger gDNA		177	84
dideoxysequencing (Sanger gDNA		154297	86
dideoxysequencing (Sanger gDNA		20768	86
dideoxysequencing (Sanger gDNA		231338	86
dideoxysequencing (Sanger gDNA		167	86
dideoxysequencing (Sanger gDNA		237804	84
dideoxysequencing (Sanger gDNA		403	84
dideoxysequencing (Sanger gDNA		132277	84
dideoxysequencing (Sanger gDNA		29629	84
dideoxysequencing (Sanger gDNA		21437	84
dideoxysequencing (Sanger gDNA		690	84
dideoxysequencing (Sanger gDNA		403	86
dideoxysequencing (Sanger gDNA		132744	86
dideoxysequencing (Sanger gDNA		132277	86
dideoxysequencing (Sanger gDNA		29629	86
dideoxysequencing (Sanger gDNA		132744	84
dideoxysequencing (Sanger gDNA		690	86
dideoxysequencing (Sanger gDNA		21437	86
dideoxysequencing (Sanger gDNA		192643	85
dideoxysequencing (Sanger gDNA		175607	85
dideoxysequencing (Sanger gDNA		585	85
dideoxysequencing (Sanger gDNA		3791	86
dideoxysequencing (Sanger gDNA		164570	86
dideoxysequencing (Sanger gDNA		152071	86
dideoxysequencing (Sanger gDNA		8929	86
dideoxysequencing (Sanger gDNA		2879	86
dideoxysequencing (Sanger gDNA		67381	1
dideoxysequencing (Sanger gDNA		74971	1
dideoxysequencing (Sanger gDNA		90905	1
dideoxysequencing (Sanger gDNA		92351	1
dideoxysequencing (Sanger gDNA		121590	41
dideoxysequencing (Sanger gDNA		61605	42
dideoxysequencing (Sanger gDNA		52959	43
dideoxysequencing (Sanger gDNA		61131	44
dideoxysequencing (Sanger gDNA		59679	45
dideoxysequencing (Sanger gDNA		50980	46
dideoxysequencing (Sanger gDNA		129655	47
dideoxysequencing (Sanger gDNA		79303	48
dideoxysequencing (Sanger gDNA		78304	49
dideoxysequencing (Sanger gDNA		124435	50
dideoxysequencing (Sanger gDNA		125440	51

dideoxysequencing (Sanger gDNA	722	51
dideoxysequencing (Sanger gDNA	138033	52
dideoxysequencing (Sanger gDNA	128885	53
dideoxysequencing (Sanger gDNA	127362	54
dideoxysequencing (Sanger gDNA	127122	55
dideoxysequencing (Sanger gDNA	257581	56
dideoxysequencing (Sanger gDNA	142743	57
dideoxysequencing (Sanger gDNA	135325	58
dideoxysequencing (Sanger gDNA	296355	59
dideoxysequencing (Sanger gDNA	131798	60
dideoxysequencing (Sanger gDNA	121662	61
dideoxysequencing (Sanger gDNA	133051	62
dideoxysequencing (Sanger gDNA	120671	63
dideoxysequencing (Sanger gDNA	101424	64
dideoxysequencing (Sanger gDNA	1284	64
dideoxysequencing (Sanger gDNA	222080	65
dideoxysequencing (Sanger gDNA	189052	66
dideoxysequencing (Sanger gDNA	131529	67
dideoxysequencing (Sanger gDNA	1124	68
dideoxysequencing (Sanger gDNA	358028	68
dideoxysequencing (Sanger gDNA	436401	69
dideoxysequencing (Sanger gDNA	147126	70
dideoxysequencing (Sanger gDNA	892	70
dideoxysequencing (Sanger gDNA	692255	71
dideoxysequencing (Sanger gDNA	134347	72
dideoxysequencing (Sanger gDNA	140814	73
dideoxysequencing (Sanger gDNA	77538	74
dideoxysequencing (Sanger gDNA	65670	75
dideoxysequencing (Sanger gDNA	741	87
dideoxysequencing (Sanger gDNA	759	2
dideoxysequencing (Sanger gDNA	736	3
dideoxysequencing (Sanger gDNA	678	4
dideoxysequencing (Sanger gDNA	699	5
dideoxysequencing (Sanger gDNA	711	6
dideoxysequencing (Sanger gDNA	678	76
dideoxysequencing (Sanger gDNA	730	77
dideoxysequencing (Sanger gDNA	626	78
dideoxysequencing (Sanger gDNA	66023	79
dideoxysequencing (Sanger gDNA	90515	80
dideoxysequencing (Sanger gDNA	92501	81
dideoxysequencing (Sanger gDNA	715	82
dideoxysequencing (Sanger gDNA	128982	83
dideoxysequencing (Sanger gDNA	1457	20
dideoxysequencing (Sanger gDNA	1445	20
dideoxysequencing (Sanger gDNA	1477	20
dideoxysequencing (Sanger gDNA	1492	20
dideoxysequencing (Sanger gDNA	1422	20
dideoxysequencing (Sanger gDNA	1432	20
dideoxysequencing (Sanger gDNA	1437	20
dideoxysequencing (Sanger gDNA	1334	20
dideoxysequencing (Sanger gDNA	47691	80

dideoxysequencing (Sanger gDNA	51787	182
dideoxysequencing (Sanger gDNA	49595	182
dideoxysequencing (Sanger gDNA	59812	183
dideoxysequencing (Sanger gDNA	99288	184
dideoxysequencing (Sanger gDNA	49597	184
dideoxysequencing (Sanger gDNA	59079	185
dideoxysequencing (Sanger gDNA	99781	186
dideoxysequencing (Sanger gDNA	52118	186
dideoxysequencing (Sanger gDNA	108982	187
dideoxysequencing (Sanger gDNA	718	187
dideoxysequencing (Sanger gDNA	329129	188
dideoxysequencing (Sanger gDNA	9953	188
dideoxysequencing (Sanger gDNA	9741	188
dideoxysequencing (Sanger gDNA	61020	189
dideoxysequencing (Sanger gDNA	60932	190
dideoxysequencing (Sanger gDNA	346952	191
dideoxysequencing (Sanger gDNA	50609	191
dideoxysequencing (Sanger gDNA	60987	192
dideoxysequencing (Sanger gDNA	46052	193
dideoxysequencing (Sanger gDNA	110720	194
dideoxysequencing (Sanger gDNA	101558	195
dideoxysequencing (Sanger gDNA	10720	195
dideoxysequencing (Sanger gDNA	50095	195
dideoxysequencing (Sanger gDNA	107966	196
dideoxysequencing (Sanger gDNA	11278	197
dideoxysequencing (Sanger gDNA	96463	197
dideoxysequencing (Sanger gDNA	11713	198
dideoxysequencing (Sanger gDNA	99271	198

21	13.2 3rd leg	13.99
3	2.1 3rd leg	1.99
1.7	1 3rd leg	1.67
0.2	1.7 3rd leg	0.18
0.16	2 3rd leg	0.1
0.13	2 3rd leg	0.08
0.14	1.7 3rd leg	0.08
0.23	1.7 3rd leg	0.13
	2 3rd leg	
0.5	1.6 4th leg	0.51
0.33	2 4th leg	0.26
0.07	2 4th leg	0.17
0.11	1.1 4th leg	0.18
0.22	2 4th leg	0.26
0.22	2 4th leg	0.26
0.4	2.2 4th leg	0.35
0.35	2 4th leg	0.33
0.4	2.1 4th leg	0.36
	19 4th leg	
	19 4th leg	
0.35	12 4th leg	0.36
	0.1 4th leg	
	0.1 4th leg	
	0.2 4th leg	
0.36	2.1 4th leg	0.33
0.28	1.7 4th leg	0.29
0.65	2.1 4th leg	0.4
0.21	1.8 5th leg	0.22
	1.8 5th leg	
	2 5th leg	
	2.2 5th leg	
	2 5th leg	
	1.7 5th leg	
	1.9 5th leg	
	2 5th leg	
	1.7 5th leg	
	1.9 5th leg	
	30 5th leg	
0.071	1.4 6th leg	
0.071	1.4 6th leg	
	1.2 6th leg	
	1 6th leg	
	2.07	
	2.07	
	2.07	
	2.07	
	2.07	
	2.07	
	2.07	
	2.07	
0.071	1.4 6th leg	

0.501	1.8
0.501	1.8
0.098	1.5
0.12	1.5
0.12	1.5
0.126	1.8
0.128	1.8
0.128	1.8
0.17	1.8
0.17	1.8
0.009	1.5
0.009	1.5
0.009	1.5
0.144	1.5
0.206	1.5
0.103	1.8
0.103	1.8
0.074	2
0.084	2.8
0.137	1.5
0.15	1.9
0.15	1.9
0.15	1.9
0.142	2.2
	0.3
	0.3
	1.5
	1.5

salinity - (psu water depth - (m temperature - (C chlorophyll density - (ug/ chlorophyll density/annual - (ug

32.4	7	25.8		0.501
32.5	4573	26.6		0.128
36.6	4200	20	0.17	
36.7	4200	20.5	0.17	
36.7	4200	20.5	0.17	
36.7	4200	20.5	0.17	
36.7	4200	20.5	0.17	
36.6	4200	20	0.17	
36.7	4200	20.5	0.17	
36.6	4200	20	0.17	
36.6	4200	20	0.17	
36.6	4200	20	0.17	
36.7	4200	20.5	0.17	
36.7	4200	20.5	0.17	
36.7	4200	20.5	0.17	
36.7	4200	20.5	0.17	
36.7	4200	20.5	0.17	
36.7	4200	20.5	0.17	
36.6	4200	20	0.17	
36.6	4200	20	0.17	
36.6	4200	20	0.17	
36.6	4200	20	0.17	
36.7	4200	20.5	0.17	
36.6	4200	20	0.17	
36.6	4200	20	0.17	
36.7	4200	19.8		0.07
36.7	4200	19.8		0.07
36.7	4200	19.8		0.07
36.6	4200	20	0.17	
36.6	4200	20	0.17	
36.6	4200	20	0.17	
36.6	4200	20	0.17	
36.6	4200	20	0.17	
36.7	4200	22.9	0.1	
36.7	4200	22.9	0.1	
36.7	4200	22.9	0.1	
36.7	4200	22.9	0.1	
29.2	106	18.2		0.93
29.9	119	11.7		0.93
28.3	142	17.3		0.61
30.2	64	15		5.78
	11	11.2		1.69
31.7	139	17.9		0.93
26.5	12	9.4		1.42
31	32	11		2.48
31	10	12		2.42
	8	11		8.21
3.5	25	1		13.99

3.5	25	1		13.99
	20	9.3		1.99
	31	18.6		1.67
36	47	25		0.18
35.8	3333	26.4		0.1
35.8	4513	27		0.08
35.4	4470	27.4		0.08
35.4	3336	27.7		0.13
0.1	4.2	28.6		
30.7	76	27.6		0.51
32.3	2431	29.3		0.26
32.6	1139	28.7		0.17
31.4	30	28.3		0.18
32.6	2386	27.8		0.26
32.6	2386	27.8		0.26
34.9	2.3	25.5		0.35
	156			0.33
34.5	12	26.2		0.36
	19	26.9		
	19	26.9		
	19.6	18.6		0.36
	1.6	25.4		
	1.6	25.4		
63.4	0.3	37.6		
	35	27.5		0.33
34.5	71	21.8		0.29
34.6	67	25.8		0.4
	3334	28		0.22
	4000	28.4		
	4000	28.6		
	4000	27.8		
35	4000	28		
	4000	27.6		
35.9	4000	27.6		
	4000	27.6		
37	4000	28.3		
35.3	4000	28.7		
37.3	2400	28.6		
35.1	34	28.9	0.095	
32.6	900	28.8	0.095	
	24	27.8		
34.2	10	27.3		
6.56		15.02		
6.56		15.02		
6.56		15.02		
6.56		15.02		
6.56		15.02		
6.56		15.02		
6.56		15.02		
6.56		15.02		
35.1	34	28.9	0.095	

32.4	7	25.8	0.108
32.4	7	25.8	0.108
32.6	4573	27.2	0.138
32.7	1220	27	0.128
32.7	1220	27	0.128
32.3	3841	26.4	0.203
32.5	4573	26.6	0.134
32.5	4573	26.6	0.134
33.3	4573	27.5	0.24
33.3	4573	27.5	0.24
33.1	3649	28.2	0.14
33.1	3649	28.2	0.14
33.1	3649	28.2	0.14
33.2	3220	27.9	0.142
33.1	2150	26.2	0.287
35.5	14	26.4	0.206
35.5	14	26.4	0.206
35.4	2995	23.8	0.078
35.6	5081	22.5	0.117
35.4	4309	23.1	0.14
35.8	4921	20.2	0.151
35.8	4921	20.2	0.151
35.8	4921	20.2	0.151
35.8	1860	20.4	0.228
	1		
	1		
	5		
	5		

dissolved oxygen - (nmol/l) fluorescence - (ug/l) transmission

Table S5 - Published Gut Viromes that were Analyzed

SRA run number	Reference
SRR029698	Reyes et al., 2010. Nature
SRR029697	Reyes et al., 2010. Nature
SRR029695	Reyes et al., 2010. Nature
SRR029693	Reyes et al., 2010. Nature
SRR029691	Reyes et al., 2010. Nature
SRR029690	Reyes et al., 2010. Nature
SRR029689	Reyes et al., 2010. Nature
SRR029688	Reyes et al., 2010. Nature
SRR029686	Reyes et al., 2010. Nature
SRR029703	Reyes et al., 2010. Nature
SRR029702	Reyes et al., 2010. Nature
SRR029701	Reyes et al., 2010. Nature
SRR029700	Reyes et al., 2010. Nature
SRR029699	Reyes et al., 2010. Nature
SRR029696	Reyes et al., 2010. Nature
SRR029694	Reyes et al., 2010. Nature
SRR029692	Reyes et al., 2010. Nature
SRR043420	Minot et al., 2011, Genome Research
SRR043421	Minot et al., 2011, Genome Research
SRR043422	Minot et al., 2011, Genome Research
SRR043423	Minot et al., 2011, Genome Research
SRR043424	Minot et al., 2011, Genome Research
SRR043425	Minot et al., 2011, Genome Research
SRR043426	Minot et al., 2011, Genome Research
SRR043427	Minot et al., 2011, Genome Research
SRR043428	Minot et al., 2011, Genome Research
SRR043429	Minot et al., 2011, Genome Research
SRR043430	Minot et al., 2011, Genome Research
SRR043431	Minot et al., 2011, Genome Research
SRR043432	Minot et al., 2011, Genome Research
SRR043433	Minot et al., 2011, Genome Research
SRR043434	Minot et al., 2011, Genome Research
SRR043435	Minot et al., 2011, Genome Research
SRR043436	Minot et al., 2011, Genome Research
SRR089800	Kim et al., 2011, Applied and Environmental Microbiology
SRR089802	Kim et al., 2011, Applied and Environmental Microbiology
SRR089803	Kim et al., 2011, Applied and Environmental Microbiology
SRR089804	Kim et al., 2011, Applied and Environmental Microbiology
SRR089805	Kim et al., 2011, Applied and Environmental Microbiology