Supplementary materials

Table S 1 Characteristics of 26 geographically distributed wastewater treatment systems.

Table S 2 The projected *z* values of taxa-area relationships for functional gene categories involved in C cycling.

Table S 3 The projected *z* values of taxa-area relationships for functional gene categories involved in N cycling.

Table S 4The projected *z* values of taxa-area relationships for functional gene categories involved in sulfur cycling.

Table S5 z values of taxa-area curves for functional gene categories involved in Organic Remediation.

Figure S1 The distance-decay relationship for activated sludge microbial communities in WWTPs.

Table S1 Summery of 26 WWTPs

WWTPs	Location			Process	Capacity	COD	TN	Ammonia	TP	Temperature	DO	PH	Conductivity	MLSS
	City	Longitude	Latitude		(10^4 m^3/d)	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(℃)	(mg/L)		(μs/cm)	(mg/L)
WC_1	Harbin	126.719742	45.815879	AO[a]	15	154	55	37	6.7	14.4	5.0	7.2	693	3764
WC_2	Harbin	126.719742	45.815879	AO	15	154	55	37	6.7	14.1	4.0	7.2	712	3359
XJH	Beijing	116.298523	40.018675	A^2O[b]	3	150	48	42	6.1	19.8	2.9	7.1	928	4231
QH_1	Beijing	116.360664	40.032872	Reversed A^2O	20	437	64	59	6.1	23.1	3.8	7.2	604	2647
QH_2	Beijing	116.360664	40.032872	A^2O	20	437	64	59	6.1	23.1	3.0	7.0	595	3741
GBD_1	Beijing	116.533785	39.906247	A^2O	25	452	58	52	6	20.7	2.4	7.0	1018	5047
GBD_4	Beijing	116.533785	39.906247	A^2O	25	452	58	52	6	20.6	2.0	7.0	1029	4136
XHM	Beijing	116.443920	39.828709	A^2O	15	443	46	36	6.8	26	2.1	6.9	1356	4083
HR_1	Beijing	116.639099	40.293668	oxidation ditch	3	412	39	32	4.2	22	2.8	7.6	985	4200
HR_2	Beijing	116.639099	40.293668	MBR[c]	3.4	412	39	32	4.2	22	2.9	7.4	974	4017
KFQ	Tianjin	117.623062	39.100492	BFD[d]	1.25	439	45	40	6.0	23	2.5	7.5	5125	4329
JZZ	Tianjin	117.201204	39.071978	A^2O	15	427	47	42	5.9	23	2.5	7.6	1223	5047
LS	Dalian	121.495056	38.861365	CAST[e]	6	243	38	37	5.4	15.3	5.5	6.5	1199	3280
QS	Dalian	121.643372	39.001043	CAST	3	258	44	39	5.9	14.1	5.0	6.6	1797	3370
LHT_1	Dalian	121.679077	38.881345	CAST	8	291	18	16	5.8	13.7	4.5	6.5	3180	3250

LHT_2	Dalian	121.679077	38.881345	CAST	8	291	18	16	5.3	13.6	4.0	6.5	3370	4022
MTG_1	Zhengzhou	113.742142	34.848748	A^2O	8	427	50	37	10.8	17	3.0	6.8	823	5608
MTG_2	Zhengzhou	113.742142	34.848748	A^2O	8	427	50	37	10.8	17	3.0	6.8	829	5817
WXZ_1	Zhengzhou	113.797760	34.779126	A^2O	6	443	55	48	5.7	17	3.0	6.9	845	3226
WXZ_2	Zhengzhou	113.797760	34.779126	A^2O	15	443	55	48	5.7	17	3.0	6.9	838	3572
LC_1	Wuxi	120.316772	31.530848	A^2O	10	374	45	45	7.2	18.4	2.3	6.9	950.5	5489
LC_2	Wuxi	120.316772	31.530848	A^2O	10	374	45	45	7.2	18.05	2.5	6.7	991	5612
CB_2	Wuxi	120.332909	31.611873	Oxidation ditch	5	324	39	34	4.3	19.6	5.6	7.1	1290	7988
CB_1	Wuxi	120.332909	31.611873	MBR	5	324	39	34	4.3	19.0	2.7	6.9	1340	8225
ZY	Shanghai	121.617966	31.343668	A^2O	25	281	58	41	6	26.8	2.8	7.8	683	3587
QY	Shanghai	121.488104	31.279504	A^2O	6	300	64	39	7.2	26.4	2.8	7.5	921	4421
SH	Wuhan	114.344330	30.560783	A^2O	5	298	45	39	7	29.1	2.6	7.2	368	3791
CSY	Changsha	113.063049	28.222508	A^2O	12	177	22	20	6	25.1	4.6	6.8	648	4500
GZ	Changsha	113.022194	28.183288	Oxidation ditch	16	174	25	25	5.8	25	3.2	7.0	336	2800
LF_1	Shenzhen	114.147949	22.542367	AB[f]	10	317	33	27	4.3	25	4.5	6.7	700	4700
LF_2	Shenzhen	114.147949	22.542367	MUCT[g]	25	328	33	26	4.3	24.6	3.0	6.8	645	4800
BH	Shenzhen	114.104347	22.534915	MSBR[h]	18	343	77.9	55	5.1	30.7	2.6	6.8	740	4600

[a] Anoxic/Aerobic

[b] Anaerobic/Anoxic/Aerobic

[c] Membrane Bioreactor

[d] Biological Fluidized Bed

[e] Cyclic Activated Sludge Technology

[f] Absorption Biodegradation

[g] Modified University of Cape Town

[h] Modified Sequencing Batch Reactor

Table S2 *z* values of taxa-area curves for functional gene categories involved in C cycling.

Gene category	z	t	n	p	95% C.I.	95% C.I.
rubisco	0.0051	-444.254	9999	<0.01	-1.75E-05	2.77E-05
xylA	0.0040	-411.224	9999	<0.01	-3.18E-05	6.14E-06
endochitinase	0.0043	-488.305	9999	<0.01	-1.27E-05	2.17E-05
CODH	0.0036	-413.297	9999	<0.01	-1.72E-05	1.69E-05
xylanase	0.0049	-419.757	9999	<0.01	-2.46E-05	2.14E-05
amyX	0.0084	-331.724	9999	<0.01	-6.57E-05	3.34E-05
cda	0.0056	-530.245	9999	<0.01	-1.37E-05	2.76E-05
exochitinase	0.0048	-457.925	9999	<0.01	-8.93E-06	3.22E-05
mcrA	0.0072	-425.019	9999	<0.01	-6.64E-05	4.82E-07
vanA	0.0027	-457.412	9999	<0.01	-1.66E-05	6.89E-06
cellobiase	0.0040	-430.567	9999	<0.01	-2.36E-05	1.29E-05
pcc	0.0034	-442.273	9999	<0.01	-1.09E-05	1.92E-05
CDH	0.0021	-411.655	9999	<0.01	-9.17E-06	1.06E-05
acetylglucosaminidase	0.0042	-504.892	9999	<0.01	-1.07E-05	2.16E-05
ara	0.0037	-439.753	9999	<0.01	-1.98E-05	1.32E-05
phenol_oxidase	0.0054	-489.25	9999	<0.01	-4.49E-05	-1.58E-06
isopullulanase	0.0007	-25.9738	9999	<0.01	-4.03E-05	6.57E-05
AceB	0.0037	-438.404	9999	<0.01	-2.93E-05	3.71E-06
amyA	0.0040	-509.384	9999	<0.01	-1.33E-05	1.77E-05
glucoamylase	0.0046	-446.542	9999	<0.01	-2.33E-05	1.75E-05
AceA	0.0041	-454.18	9999	<0.01	-1.69E-05	1.87E-05
camDCAB	0.0024	-140.793	9999	<0.01	-4.54E-05	2.14E-05
pulA	0.0043	-451.771	9999	<0.01	-3.47E-05	2.39E-06
glx	0.0041	-474.507	9999	<0.01	-2.39E-05	9.76E-06
lip	0.0024	-350.574	9999	<0.01	-1.17E-05	1.51E-05
nplT	0.0062	-512.912	9999	<0.01	-1.89E-05	2.84E-05
exoglucanase	0.0047	-446.94	9999	<0.01	-1.13E-05	3.03E-05
endoglucanase	0.0037	-472.43	9999	<0.01	-1.32E-05	1.72E-05
pectinase	0.0048	-385.568	9999	<0.01	-1.63E-05	3.26E-05
aclB	0.0056	-413.935	9999	<0.01	-3.64E-05	1.71E-05
ara_fungi	0.0043	-465.398	9999	<0.01	-1.60E-05	2.03E-05
FTHFS	0.0051	-399.154	9999	<0.01	-1.80E-05	3.17E-05
pmoA	0.0068	-463.087	9999	<0.01	-3.37E-05	2.41E-05
vdh	0.0029	-276.636	9999	<0.01	-2.61E-05	1.48E-05
mannanase	0.0038	-426.718	9999	<0.01	-2.02E-05	1.43E-05
mnp	0.0043	-375.742	9999	<0.01	-1.87E-05	2.57E-05
mmoX	0.0057	-360.141	9999	<0.01	-1.83E-05	4.37E-05
limEH	0.0039	-366.052	9999	<0.01	-2.13E-05	2.05E-05
AssA	0.0037	-236.654	9999	<0.01	-3.68E-05	2.50E-05
apu	0.0033	-181.365	9999	<0.01	-1.60E-05	5.47E-05
LMO	0.0056	-266.673	9999	<0.01	-2.80E-05	5.43E-05

Table S3 *z* values of taxa-area curves for functional gene categories involved in N cycling.

Gene category	*z* value	t	n	p	95% C.I.	95% C.I.
nasA	0.0035	-404.222	9999	<0.01	-9.26E-06	2.42E-05
norB	0.0036	-399.995	9999	<0.01	-1.59E-05	1.92E-05
ureC	0.0036	-478.16	9999	<0.01	-9.11E-06	2.01E-05
napA	0.0055	-436.097	9999	<0.01	-2.22E-05	2.76E-05
narG	0.0028	-401.127	9999	<0.01	-1.19E-05	1.55E-05
nifH	0.0057	-501.851	9999	<0.01	-1.27E-05	3.16E-05
nosZ	0.0048	-423.44	9999	<0.01	-3.53E-05	9.80E-06
NiR	0.0036	-422.542	9999	<0.01	-1.66E-05	1.71E-05
nirS	0.0048	-447.916	9999	<0.01	-3.99E-05	1.99E-06
nrfA	0.0048	-440.518	9999	<0.01	-1.42E-05	2.82E-05
gdh	0.0065	-427.784	9999	<0.01	-3.65E-05	2.30E-05
nirK	0.0046	-448.159	9999	<0.01	-1.52E-05	2.53E-05
nirA	0.0034	-309.608	9999	<0.01	-2.01E-05	2.28E-05
hao	0.0064	-377.463	9999	<0.01	-2.35E-05	4.31E-05
NirB	0.0024	-240.02	9999	<0.01	-1.55E-05	2.37E-05
hzo	0.0125	-248.349	9999	<0.01	-0.000145918	5.18E-05

Table S4 *z* values of taxa-area curves for functional gene categories involved in sulfur cycling.

Gene category	*z* value	t	n	p	95% C.I.	95% C.I.
dsrB	0.0051	-468.797	9999	<0.01	-1.93E-05	2.36E-05
dsrA	0.0060	-470.104	9999	<0.01	-4.15E-05	8.75E-06
APS_AprA	0.0068	-481.016	9999	<0.01	-1.37E-05	4.15E-05
sox	0.0033	-424.335	9999	<0.01	-1.09E-05	1.93E-05
CysJ	0.0039	-455.765	9999	<0.01	-2.09E-05	1.29E-05
APS_AprB	0.0057	-418.506	9999	<0.01	-1.97E-05	3.37E-05
sir	0.0047	-447.837	9999	<0.01	-2.68E-05	1.44E-05
sqr	0.0031	-243.896	9999	<0.01	-1.82E-05	3.12E-05
fccAB	0.0041	-439.801	9999	<0.01	-1.28E-05	2.40E-05
AprA	0.0045	-421.325	9999	<0.01	-2.52E-05	1.65E-05
CysI	0.0040	-248.044	9999	<0.01	-7.85E-05	-1.38E-05

Table S5 *z* values of taxa-area curves for functional gene categories involved in Organic Remediation.

Gene category	z value	t	n	p	95% C.I.	95% C.I.
mdlC	0.0025	-424.765	9999	<0.01	-2.20E-06	2.09E-05
akbF	0.0036	-407.204	9999	<0.01	-7.16E-06	2.78E-05
mauAB	0.0017	-286.584	9999	<0.01	-1.19E-05	1.15E-05
pcpB	0.0023	-398.011	9999	<0.01	-9.48E-06	1.30E-05
catB	0.0042	-477.821	9999	<0.01	-1.20E-05	2.21E-05
exaA	0.0031	-457.333	9999	<0.01	-1.53E-05	1.12E-05
dmsA	0.0045	-459.121	9999	<0.01	-1.66E-05	2.18E-05
hmgC	0.0033	-398.847	9999	<0.01	-1.79E-05	1.49E-05
proO	0.0032	-386.86	9999	<0.01	-4.68E-06	2.80E-05
pimF	0.0033	-422.407	9999	<0.01	-1.63E-05	1.46E-05
nmoA	0.0036	-447.359	9999	<0.01	-1.00E-06	3.03E-05
atzA	0.0052	-505.274	9999	<0.01	-3.45E-05	6.23E-06
HcaB	0.0029	-310.368	9999	<0.01	-1.85E-05	1.80E-05
atzB	0.0035	-409.289	9999	<0.01	-2.20E-05	1.14E-05
pcaG	0.0024	-368.732	9999	<0.01	-1.79E-05	7.51E-06
PobA	0.0029	-403.66	9999	<0.01	-1.89E-05	9.46E-06
nbaC	0.0034	-351.028	9999	<0.01	-2.47E-05	1.30E-05
ChnA	0.0059	-363.58	9999	<0.01	-3.68E-05	2.69E-05
adpB	0.0039	-308.426	9999	<0.01	-4.14E-05	8.07E-06
PhaB	0.0043	-411.796	9999	<0.01	-2.86E-05	1.26E-05
mdlA	0.0037	-420.953	9999	<0.01	-3.11E-05	3.56E-06
bclA	0.0030	-398.374	9999	<0.01	-1.30E-05	1.65E-05
pobA	0.0023	-378.438	9999	<0.01	-7.57E-06	1.60E-05
nagI	0.0033	-420.235	9999	<0.01	-2.21E-05	9.07E-06
nagG	0.0036	-474.752	9999	<0.01	-2.16E-05	8.56E-06
arhA	0.0022	-393.402	9999	<0.01	-1.17E-05	1.02E-05
nicdehydr	0.0034	-276.222	9999	<0.01	-4.40E-05	4.35E-06
BADH	0.0035	-395.836	9999	<0.01	-2.80E-05	6.74E-06
Catechol	0.0023	-392.533	9999	<0.01	-1.17E-05	1.13E-05
tfdA	0.0030	-412.507	9999	<0.01	-1.55E-05	1.34E-05
ChnE	0.0025	-194.47	9999	<0.01	-3.10E-05	1.96E-05
Arylest	0.0042	-489.183	9999	<0.01	-2.86E-05	4.83E-06
alkB	0.0031	-421.963	9999	<0.01	-2.38E-05	4.97E-06
phn	0.0032	-443.562	9999	<0.01	-2.39E-05	4.65E-06
alkH	0.0036	-351.592	9999	<0.01	-2.76E-05	1.25E-05
hmgB	0.0021	-328.225	9999	<0.01	-2.18E-05	3.62E-06
linB	0.0040	-454.186	9999	<0.01	-1.41E-05	2.01E-05
nbzC	0.0009	-66.922	9999	<0.01	-2.21E-05	3.17E-05
tutFDG	0.0058	-450.706	9999	<0.01	-2.84E-06	4.75E-05
hmgA	0.0031	-444.677	9999	<0.01	-1.47E-05	1.25E-05
dehH	0.0050	-417.716	9999	<0.01	-2.16E-05	2.56E-05
benD	0.0042	-403.13	9999	<0.01	-2.61E-05	1.50E-05
GCoADH	0.0040	-397.208	9999	<0.01	-2.79E-05	1.20E-05
bphC	0.0027	-424.191	9999	<0.01	-1.25E-05	1.23E-05
xylG	0.0027	-387.053	9999	<0.01	-2.31E-05	4.88E-06
alkK	0.0053	-456.432	9999	<0.01	-2.40E-05	2.16E-05

pheA	0.0038	-436.995	9999	<0.01	-5.16E-06	2.86E-05
phtA	0.0037	-471.74	9999	<0.01	-2.92E-05	2.12E-06
chnB	0.0031	-405.55	9999	<0.01	-2.31E-05	6.61E-06
mhpA	0.0030	-373.278	9999	<0.01	-3.02E-06	2.84E-05
tftH	0.0025	-350.835	9999	<0.01	-1.36E-05	1.48E-05
catechol_B	0.0037	-396.08	9999	<0.01	-2.28E-05	1.35E-05
phdCI	0.0040	-397.06	9999	<0.01	-3.00E-05	9.32E-06
atzC	0.0025	-394.513	9999	<0.01	-8.35E-06	1.64E-05
nitA	0.0033	-406.413	9999	<0.01	-1.25E-05	1.95E-05
nagK	0.0040	-454.652	9999	<0.01	-6.56E-06	2.77E-05
THL	0.0023	-19.3882	9999	<0.01	-0.000543196	3.77E-05
xylF	0.0040	-390.366	9999	<0.01	-7.83E-07	3.95E-05
trzN	0.0022	-327.491	9999	<0.01	-9.24E-07	2.56E-05
bbs	0.0004	-21.1725	9999	<0.01	-2.28E-05	4.33E-05
linC	0.0033	-406.266	9999	<0.01	-2.77E-05	4.69E-06
chnC	0.0062	-146.155	9999	<0.01	-8.46E-05	8.24E-05
PceA	0.0074	-328.568	9999	<0.01	-1.05E-05	7.70E-05
Quinoline	0.0031	-163.153	9999	<0.01	-1.37E-05	6.14E-05
nhh	0.0031	-402.04	9999	<0.01	-1.69E-05	1.37E-05
pcpE	0.0034	-382.22	9999	<0.01	-1.32E-05	2.21E-05
tfdB	0.0034	-346.399	9999	<0.01	-1.36E-05	2.51E-05
benAB	0.0016	-338.818	9999	<0.01	-8.24E-06	1.06E-05
cbdA	0.0027	-367.203	9999	<0.01	-1.30E-05	1.59E-05
xylC	0.0022	-301.52	9999	<0.01	-1.04E-05	1.84E-05
nahA	0.0031	-474.14	9999	<0.01	-1.45E-05	1.09E-05
mdlD	0.0031	-417.677	9999	<0.01	-1.57E-05	1.33E-05
Xamo	0.0033	-335.874	9999	<0.01	-1.01E-05	2.78E-05
scnABC	0.0026	-367.726	9999	<0.01	-2.17E-05	5.73E-06
alkJ	0.0064	-422.479	9999	<0.01	-2.41E-05	3.55E-05
pnbA	0.0027	-285.88	9999	<0.01	-1.19E-06	3.61E-05
carA	0.0026	-382.763	9999	<0.01	-1.56E-05	1.12E-05
bco	0.0040	-420.015	9999	<0.01	-3.02E-05	7.13E-06
rd	0.0062	-445.701	9999	<0.01	-2.73E-05	2.77E-05
dehH109	0.0033	-374.611	9999	<0.01	-2.31E-05	1.14E-05
tomA	0.0003	-32.8879	9999	<0.01	-7.62E-06	2.20E-05
nitro	0.0037	-375.831	9999	<0.01	-2.72E-05	1.20E-05
phdI	0.0014	-102.213	9999	<0.01	-1.01E-05	4.16E-05
sdsA	0.0023	-117.904	9999	<0.01	-5.37E-05	2.39E-05
bphA	0.0031	-469.947	9999	<0.01	-6.40E-06	1.95E-05
nahB	0.0063	-37.7013	9999	<0.01	-0.000235619	0.00041
trzA	0.0034	-204.812	9999	<0.01	-4.45E-05	2.11E-05
oxdB	0.0031	-359.623	9999	<0.01	-3.01E-05	4.12E-06
ohbAB	0.0036	-394.406	9999	<0.01	-1.22E-05	2.31E-05
GcdB	0.0047	-500.906	9999	<0.01	-2.71E-05	9.71E-06
badK	0.0102	-37.299	9999	<0.01	-0.001117052	1.53E-05
ophC	0.0069	-405.992	9999	<0.01	-3.62E-05	3.08E-05
hdnO	0.0036	-331.957	9999	<0.01	-3.43E-05	8.96E-06
bbsG	-0.0002	31.53905	9999	<0.01	-6.13E-06	1.83E-05
bphB	0.0034	-102.277	9999	<0.01	-6.16E-05	6.73E-05
trzE	0.0026	-259.151	9999	<0.01	-1.93E-05	1.99E-05
nagL	0.0010	-160.714	9999	<0.01	-1.38E-05	9.66E-06

cumA	0.0033	-361.199	9999	<0.01	-2.82E-05	7.73E-06
atzD	0.0034	-315.53	9999	<0.01	-2.15E-05	2.13E-05
todC	0.0042	-352.767	9999	<0.01	-7.41E-06	3.93E-05
thmAB	0.0040	-165.606	9999	<0.01	-4.25E-05	5.20E-05
xylJ	0.0016	-217.082	9999	<0.01	-1.13E-05	1.74E-05
fcbA	0.0037	-338.387	9999	<0.01	-1.90E-05	2.40E-05
cmuA	0.0059	-398.97	9999	<0.01	-1.88E-05	3.89E-05
ALN	0.0034	-267.179	9999	<0.01	-5.49E-05	4.89E-05
bphD	0.0037	-349.932	9999	<0.01	-2.67E-05	1.47E-05
mdlB	0.0053	-344.005	9999	<0.01	-3.69E-05	2.39E-05
phdJ	0.0079	-58.4995	9999	<0.01	-0.0001904	0.000332
AmiE	0.0032	-322.205	9999	<0.01	-5.45E-08	3.88E-05
cymA	0.0516	-174.43	9999	<0.01	-0.000923623	0.000246
nahF	0.0073	-267.547	9999	<0.01	-6.52E-05	4.20E-05
pcpA	0.0057	-229.519	9999	<0.01	-2.81E-05	6.98E-05
nbz	-0.0017	17.02326	9999	<0.01	-0.00016263	0.000249
pchCF	0.0031	-253.101	9999	<0.01	-2.31E-05	2.47E-05
xylXY	0.0030	-211.934	9999	<0.01	-4.82E-05	7.50E-06
xylL	0.0016	-147.318	9999	<0.01	-2.69E-05	1.70E-05
POBMO	-0.0003	3.129575	9999	<0.01	-0.000223006	3.29E-05
flnB	-0.0032	17.14831	9999	<0.01	-0.000489065	0.000197
mhpC	0.0045	-407.297	9999	<0.01	-1.40E-05	2.97E-05
linD	0.0048	-26.9509	9999	<0.01	-0.000166388	0.000505
BpH	0.0091	-433.502	9999	<0.01	-2.84E-05	5.37E-05
xlnD	0.0019	-203.156	9999	<0.01	9.76E-06	4.59E-05
cmtAb	0.0012	-147.782	9999	<0.01	2.29E-07	3.14E-05
tdnB	0.0014	-86.9365	9999	<0.01	-3.97E-05	2.50E-05
CDD	0.0081	-232.984	9999	<0.01	-3.25E-05	0.000104

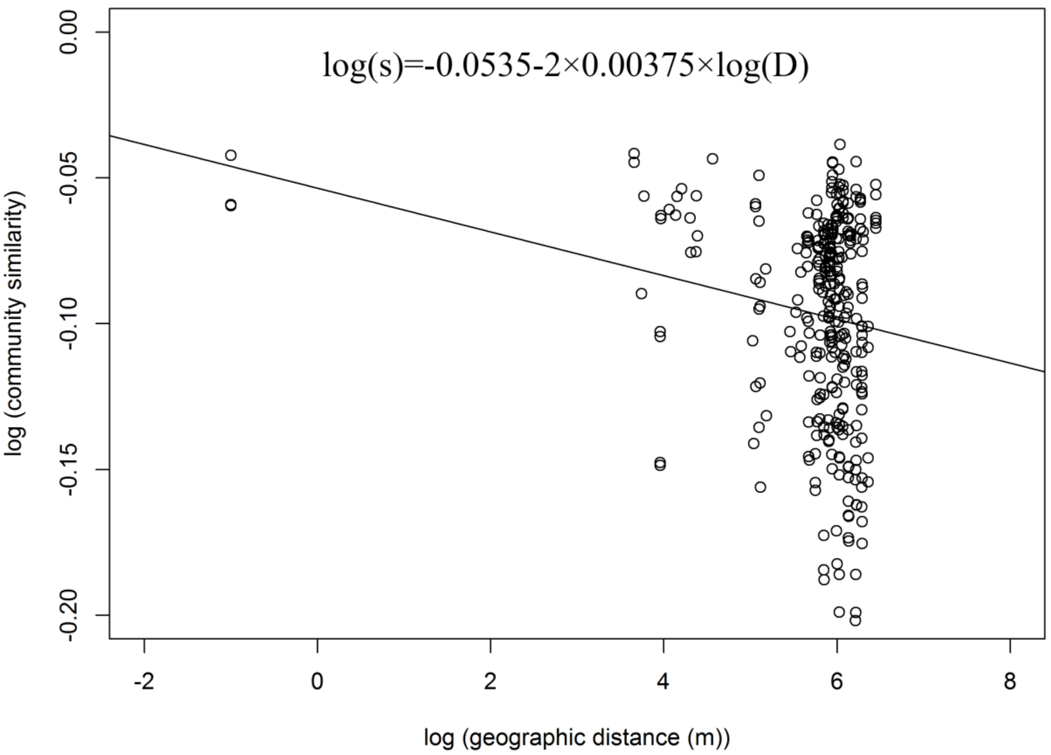

Figure S1 The distance-decay relationship for activated sludge microbial communities in WWTPs.