List of supplemental information

1. Statistical analyses

Firstly, the correlation matrix of ASV was (with relative abundance > 0.05%) calculated, and based on the correlation matrix, the UPGMA method was used for hierarchical clustering analysis. Next, the reliability of the clustering topology was evaluated using the bootstrap method. A random resampling of the ASV table was done, again calculated the correlation coefficient, did clustering and repeated the procedure 1000 times. Having thus obtained 1000 results of hierarchical clustering. The occurrence of clusters was estimated in the initial clustering among 1000 clustering results obtained from resamplings. And to identify clusters with high bootstrap support (which were found in 90% of 1000 clustering results). The visualization was a heat map with clustering using ggplot2 (heatmap.2) in R packages.

2. The code of network analysis

```
library(psych)
rm(list = ls())
otu <- read.csv("", header=T, row.names=1,    comment.char="")
otu<-t(otu)
otu <- otu[,colSums(otu)/sum(otu)>=(0.05/100)]
otu<-log(data,10)
otu_tax <- read.csv("", header=T, row.names= 1,    comment.char="")
design <- read.csv("", header=T, row.names= 1, )
otu_tax<-as.data.frame(otu_tax[colnames(otu),])
```

```
otu_abundance <- colSums(otu)

otu_pro <- cbind(otu_tax,otu_abundance)

occor = corr.test(otu,use="pairwise",method="spearman",adjust="fdr",alpha=0.05)

occor.r = occor$r

occor.p = occor$p

occor.r[occor.p>0.001|abs(occor.r)<0.9] = 0

write.csv(occor.r,file="")
```

Table S1. Estimates of richness and diversity for amplicon sequence variants (ASVs) for 12 water samples obtained from coal mining subsidence lakes and its connected rivers.

	Samples	ASV	Chao 1	Shannon	Coverage	Reads
GM	L1	2824	3514.62	7.389	0.978	63856
	L2	2945	3550.77	7.372	0.990	61963
	L3	2839	3423.15	7.265	0.979	85254
SL	L4	2833	3511.10	7.451	0.979	64080
	L5	3294	3694.01	7.575	0.973	62381
	L6	2658	3232.85	7.284	0.981	67764
	L7	2755	3301.72	7.325	0.978	75772
	L8	2824	3523.36	7.318	0.981	59969
HR	L9	1555	1815.74	7.082	0.990	94037
	L10	1807	2152.68	7.375	0.987	99111
	L11	979	1084.19	7.156	0.996	98755
	L12	1183	1398.63	6.303	0.993	105697

Table S2. Physical and chemical indices of water samples in different sites in coal mining subsidence lakes and its connected rivers.

	Samples	pH	TDS	EC	DO	T	SD	COD_{cr}	TP	TN	NH_4^+	PO_4^{3-}	SO_4^{2-}	NO_3^-	F^-	CL^-	K^+	Ca^{2+}	Na^+	Mg^{2+}
			ppm	us/cm	mg/L	℃	m	mg/L	mg/L	mg/L	mg/L	mg/L	mg/L	mg/L	mg/L	mg/L	mg/L	mg/L	mg/L	mg/L
GM	L1	8.2	1004	2048	10.3	4.9	1	33	0.268	2.42	0.21	0.08	136	4.88	0.356	50.5	7.28	85.1	243	32.9
	L2	8.08	1015	2046	10.7	5	0.9	50	0.264	2.71	0.28	0.10	172	6.84	0.403	65.2	8.62	106	300	40.1
	L3	8.33	1048	2011	9.9	5.3	0.8	20	0.126	2.4	0.19	0.06	136	5.66	0.402	49.8	7.04	83.7	239	32.4
SL	L4	8.58	828	1648	10.1	5.9	1.9	12	0.048	1.22	0.06	0.04	179	5.1	0.656	71.8	9.97	113	313	40.6
	L5	7.63	810	1582	11.1	4.8	1.34	19	0.052	1.1	0.11	0.04	183	7.71	0.324	69.3	10	112	314	41.1
	L6	8.33	819	1639	11.8	5	1.64	12	0.055	0.87	0.09	0.04	135	4.72	0.318	50.2	7.14	84.4	241	32.7
	L7	8.59	826	1652	10.5	4.8	1.2	11	0.055	1.27	0.08	0.04	137	4.94	0.38	50.8	7.7	83	240	32.4
	L8	8.57	827	1651	11.3	5.1	1.5	12	0.045	1.39	0.12	0.04	39.8	14.7	0.223	42.9	6.41	76.2	143	21.6
HR	L9	8.62	1098	2193	10.6	5.9	1.08	20	0.097	1.82	0.03	0.05	174	5.94	0.232	66.7	9.5	112	310	40.0
	L10	8.63	742	1482	10.5	6	1.56	13	0.097	1.98	0.04	0.09	136	4.67	0.436	49.3	7.09	82.8	239	32.6
	L11	8.5	1103	2230	12.1	5.4	1.1	44	0.097	2.29	0.13	0.05	158	6.77	0.345	61.9	7.62	91.5	304	40.4
	L12	8.69	1073	2146	12	6	1.37	26	0.087	1.9	0.07	0.05	159	8.59	0.188	62.1	7.87	88.7	302	39.7

Table S3. Distribution of sulfur metabolism genes in each group of samples.

		Assimilatory sulfate reduction							Sulfate transport				Dissimilatory sulfate reduction and oxidation			Sulfur oxidation		
		CysJ*	**cysI**	**CysH****	**Sir****	**cysN**	**cysD**	**cysC**	**cysA**	**cysU**	**cysW**	**cysP**	**aprA**	**aprB**	**Sat***	**sseA****	**sorA**	**glpE****
GM	L1	204.4	203.6	280.0	9.5	357.6	369.3	268.9	216.2	213.5	214.6	368.4	1.3	1.3	6.3	385.9	2.3	170.5
	L2	83.7	159.1	534.7	348.1	152.2	273.7	317.4	375.0	371.1	371.1	421.9	1.1	1.1	290.7	548.5	13.2	4.7
	L3	204.2	202.9	282.3	9.3	356.3	367.7	263.8	217.2	214.7	216.1	369.4	1.3	1.3	7.5	388.2	2.9	167.3
SL	L4	213.3	208.3	282.7	5.9	359.9	371.2	263.5	219.2	216.7	217.8	382.1	1.4	1.4	4.4	394.8	3.6	174.1
	L5	210.0	204.7	272.3	7.4	356.8	367.1	266.4	212.6	210.5	211.7	369.5	1.4	1.4	4.4	379.6	2.5	170.6
	L6	213.1	209.1	279.9	4.3	356.9	364.3	266.4	216.2	213.7	214.5	375.1	1.4	1.4	2.8	393.2	2.5	176.3
	L7	205.9	201.2	274.6	7.4	354.6	365.2	261.6	212.9	210.6	211.9	366.7	1.6	1.6	3.9	381.6	2.4	170.2
	L8	212.8	206.1	283.5	5.4	360.8	368.6	264.6	219.7	216.3	217.1	374.8	1.1	1.2	2.8	391.8	1.9	174.5
HR	L9	162.4	349.9	563.4	171.5	417.4	520.3	108.9	190.1	188.7	188.7	242.1	2.1	2.1	58.9	468.3	22.3	3.5
	L10	101.2	175.0	577.7	380.7	145.6	230.6	376.5	522.6	450.4	450.4	506.3	1.3	1.3	361.4	662.2	10.6	14.1
	L11	131.9	224.7	561.7	330.1	226.1	362.1	278.9	349.8	343.6	343.6	391.8	2.2	2.2	228.4	555.8	9.6	5.1
	L12	117.8	192.1	558.8	351.5	211.3	332.5	295.3	353.5	343.8	343.8	389.1	1.2	1.2	251.9	554.6	8.5	5.4

Table S4. Topological properties of bacterial community co-occurrence networks and their Erdős-Rényi random networks in different habitats.

	Empirical network										Random network		
	N	E		Modularity	avgCC	APL	ND	AD	GD	σ	Modularity	avgCC	APL
		Positive	Negative										
SL, GM	197	732 (59.43%)	500 (40.57%)	9.93	0.51	3.05	6	12.88	0.07	3.4	0.21	0.10	2.02
HR	205	1405 (66.45%)	710 (33.55%)	2.11	0.99	1	1	20.32	0.10	34.1	0.15	0.07	2.32

SL: Subsidence Lake; GM: Gangue Mountain; HR: Huihe River; N: Nodes; E: Edges; avgCC: Average clustering coefficient; APL: Average path length; ND: network diameter; AD: Average degree; GD: Graph density; σ: small-word coefficient.

$\sigma = (avgCC/avgCC_r)/(APL/APL_r)$ and $\sigma > 1$ indicates "small-world" properties. Subscript *r* indicates the properties of the random network.

Figure S1. Rarefaction curves of richness across water samples in coal mining subsidence lakes and its connected rivers.

Figure S2. Spearman's correlations between concentration of each environmental factors and bacterial relative abundance at genus level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure S3. Cluster analysis of frequent ASVs (relative abundance > 0.05%) correlation in subsidence lake samples.

Figure S4. Cluster analysis of frequent ASVs (relative abundance > 0.05%) correlation in Huihe river samples.

Figure S1.

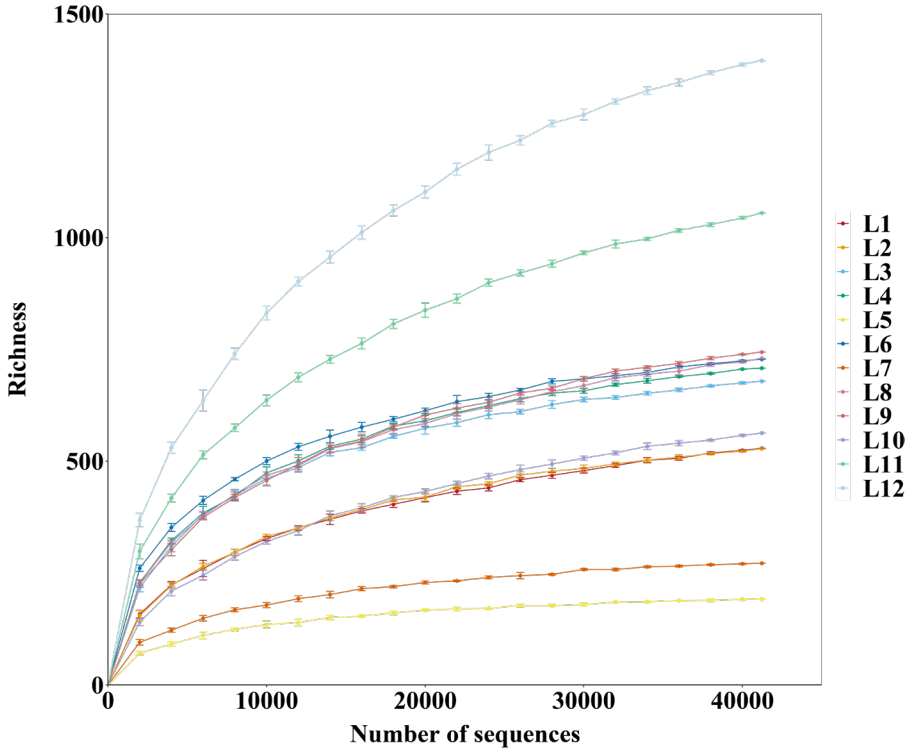
1500
1000
500
0
Richness
0
10000
20000
30000
40000
Number of sequences
L1
L2
L3
L4
L5
L6
L7
L8
L9
L10
L11
L12

Figure S2.

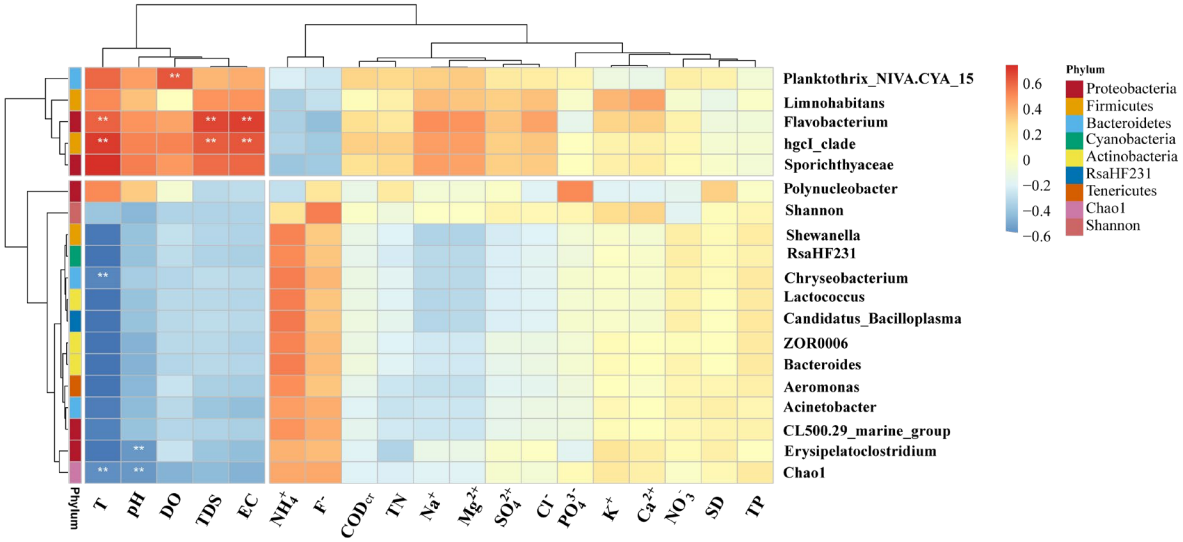
Planktothrix_NIVA.CYA_15
Limnohabitans
Flavobacterium
hgcI_clade
Sporichthyaceae
Polynucleobacter
Shannon
Shewanella
RsaHF231
Chryseobacterium
Lactococcus
Candidatus_Bacilloplasma
ZOR0006
Bacteroides
Aeromonas
Acinetobacter
CL500.29_marine_group
Erysipelatoclostridium
Chao1
Phylum
T
pH
DO
TDS
EC
NH4+
F-
CODCr
TN
Na+
Mg2+
SO42-
Cl-
PO43-
K+
Ca2+
NO3-
SD
TP
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
Phylum
Proteobacteria
Firmicutes
Bacteroidetes
Cyanobacteria
Actinobacteria
RsaHF231
Tenericutes
Chao1
Shannon

Figure S3.

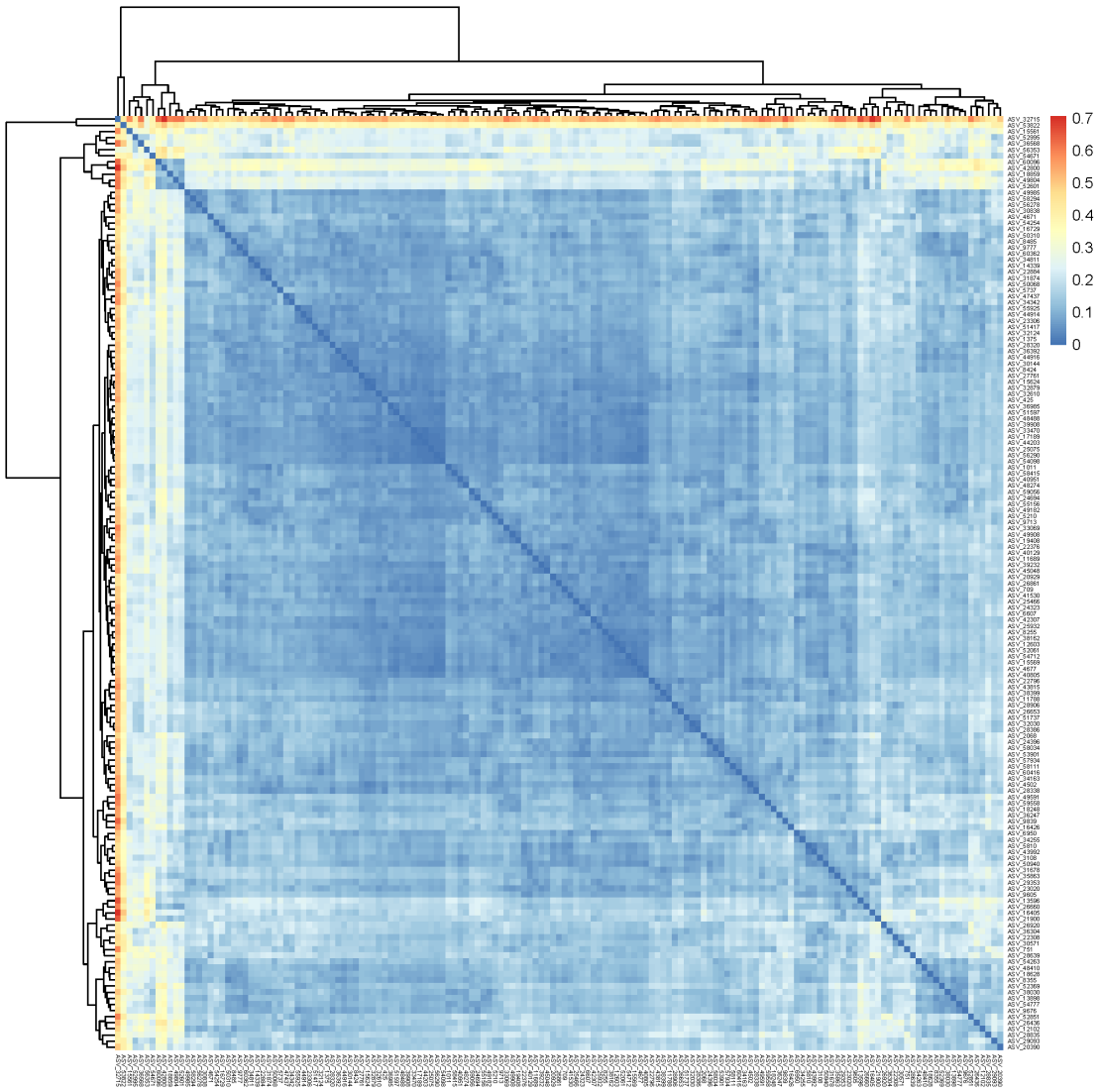

Figure S4.