



## Supplement of

## Characteristics of bacterial community in cloud water at Mt Tai: similarity and disparity under polluted and non-polluted cloud episodes

Min Wei et al.

Correspondence to: Jianmin Chen (jmchen@sdu.edu.cn)

The copyright of individual parts of the supplement might differ from the CC-BY 3.0 licence.



**Figure S1** Concentration of major ions in cloud water and atmospheric  $PM_{2.5}$  in the seven cloud episodes. Significant positive correlation was observed between  $PM_{2.5}$  (y) and major ions (x) (y=0.00477x+5.324, p<0.01, R<sup>2</sup>=0.757). High ions concentration was found in polluted cloud water samples. Cloud episodes under high atmospheric  $PM_{2.5}$  concentration and high concentration of ions were defined as polluted. After adjustment CE1-3 and CE7-2 were identified as non-polluted samples and others were polluted samples.



**Figure S2** Summary of diversity curves calculated at 97% sequence similarity. Rarefaction curves of observed OTUs continued to rise with increasing numbers of sequences, suggesting further sequencing will yield more species. However, the average Good's coverage of 13 samples was 97.2% (Table 2), indicating a comprehensive sampling of the dominant microbial groups. Moreover, the Shannon-wiener and species accumulation curves reached plateau indicating a sufficient sequencing. For the Rank-abundance curves, the wide horizontal range and smooth curves reflect the rich abundance and even species distribution. The richness estimators Chao1 predicted 1491-1999 OTUs. Chao1 estimator for the polluted samples (1676) was similar to the non-polluted samples (1680). Diversity estimators Shannon and Simpson indexes fluctuated between polluted and non-polluted samples. Bacterial diversity was higher in non-polluted samples (polluted, 3.99; non-polluted, 4.54).



**Figure S3** Hierarchical cluster (Hcluster) and principal coordinate analysis (PCoA) based on the OTUs categories. The sample similarity tree was calculated using the neighbor-joining method and the relationship among samples was determined by Bray-Curtis distance and the complete clustering method.



**Figure S4** Principal component analysis (PCA) shows the bacterial community variability between polluted and non-polluted cloud episodes. Samples in the same group indicate the close relation. The PCA plots were constructed based on Bray-Curtis similarity index calculated with the abundance of OTUs using the BIODIVERSITYR package in R (Kindt and Coe 2005). The two axes explain 78.8% of the variability for bacterial community structure. Community disparity between polluted and non-polluted samples is significant (p<0.01).



**Figure S5** Hierarchically clustered heatmap of the predominant bacterial genera distribution under polluted and non-polluted cloud episodes. Polluted cloud water samples are indicated by red square, non-polluted samples are green. The heatmap plot depicts the relative percentage of the predominant genera (variables clustering on the horizon-axis) within each sample (vertical-axis clustering). The bacterial phylogenetic tree was calculated using the neighbor-joining method and the relationship among samples was determined by Bray-Curtis distance and the complete clustering method. The relative abundance values for bacteria are indicated by color intensity with the legend indicated at the top of the figure.



**Figure S6** Bacterial taxa are related to KEGG functional pathways. Bacterial gene functions were predicted from 16S rRNA gene-based microbial compositions using the PICRUSt algorithm to make inferences from KEGG annotated databases.

	NAME	AX1	AX2	AX3	AX4
Ν					
	FR EXTRACTED	0.0585	0.1510	0.0800	0.0692
1	<b>PM</b> <sub>2.5</sub>	-0.3674	0.3099	-0.2588	0.1637
2	Τ	0.0394	-0.4939	-0.0218	0.3703
3	RH	0.0429	-0.4225	0.1192	-0.3074
4	WS	-0.0528	-0.5089	0.5716	-0.1433
5	pН	0.2176	-0.1058	-0.1572	0.4499
6	EC	-0.1772	-0.3267	-0.3409	-0.0231
7	Ions	-0.4358	0.4010	-0.1054	0.0513

Table S1 CorE: Inter set correlations of environmental variables with axes

Ν	NAME	AX1	AX2	AX3	AX4	VAR(y)	EXPL %
	FR FITTED	0.2632	0.1057	0.0905	0.0484		
1	Haliscomenobacter	0.2874	0.2901	0.4479	0.4522	0.3	47.54
2	Paracoccus	0.1034	0.3704	0.496	0.7312	1.44	78.94
3	Aquabacterium	0.3193	0.4208	0.5356	0.5486	0.24	75.47
4	Novosphingobium	0.0132	0.0855	0.3797	0.3804	0.2	41.08
5	Bdellovibrio	0.2534	0.2581	0.2942	0.3045	0.19	36.8
6	Psychrobacter	0.1277	0.2316	0.2525	0.2542	0.25	27.92
7	Pseudoalteromonas	0.6432	0.6453	0.6485	0.6487	0.23	67.71
8	Deinococcus	0.1303	0.2242	0.3175	0.3344	2.42	38.89
9	Rhodococcus	0.614	0.6143	0.6166	0.6316	0.2	77.36
10	Dietzia	0.3264	0.4314	0.4345	0.4774	0.29	52.15
11	Corynebacterium	0.7336	0.7382	0.9048	0.9063	0.22	93.83
12	Staphylococcus	0.14	0.2848	0.6314	0.6928	3.18	69.79
13	Microcoleus	0.2985	0.3061	0.3112	0.3293	0.54	45.66
14	Mesorhizobium	0.3987	0.4397	0.4623	0.4634	0.48	49.07
15	Enterobacter	0.0168	0.0266	0.1221	0.131	0.25	18.23
16	Methylobacterium	0.0487	0.1851	0.4605	0.4674	0.9	47.86
17	Caulobacter	0.089	0.147	0.3746	0.3774	0.39	57.88
18	Brevundimonas	0.0034	0.2936	0.2942	0.3737	1.59	52.51
19	Hydrotalea	0.3801	0.5155	0.5292	0.5585	0.77	56.32
20	Massilia	0.3389	0.3931	0.4012	0.4347	2.06	49.78
21	Chryseobacterium	0.308	0.4584	0.5001	0.5049	0.33	63.29
22	Bacillus	0.4788	0.4791	0.4813	0.4917	0.85	56.41
23	Pelomonas	0.3034	0.5514	0.5524	0.5749	1	58.11
24	Delftia	0.0531	0.1088	0.223	0.2249	0.74	35.26
25	Phyllobacterium	0.4656	0.4981	0.5496	0.6618	0.96	72.49
26	Empedobacter	0.3171	0.5017	0.5884	0.5891	2.96	59.36
27	Sphingomonas	0.1827	0.1829	0.1835	0.1853	1.75	23.5
28	Pseudomonas	0.1524	0.2141	0.2734	0.3872	1.54	48.71
29	Stenotrophomonas	0.0008	0.2114	0.3904	0.5081	0.63	65.78
30	Acinetobacter	0.6455	0.6456	0.6474	0.7188	3.08	72.28

Table S2 CFit: Cumulative fit per species as fraction of variance of species