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

Min Wei [1], Caihong Xu [1], Jianmin Chen [1,2,*], Chao Zhu [1], Jiarong Li [1], Ganglin Lv [1]
[1] Environment Research Institute, School of 5 Environmental Science and Engineering, Shandong University, Ji'nan 250100, China
[2] Shanghai Key Laboratory of Atmospheric Particle Pollution and Prevention (LAP), Fudan Tyndall Centre, Department of Environmental Science & Engineering, Fudan University, Shanghai 200433, China
Correspondence to: JM.Chen (jmchen@sdu.edu.cn)

Response to editor

We thank the editor for the opportunity to respond to reviewer comments. We also thank the reviewers for the beneficial comments on our manuscript. We respond to the reviewer comments in detail below. The responses to reviewer are in red. We also attach a revised manuscript with tracked changes and the amendments were highlighted with yellow color in the revised manuscript.

Response to reviewer 1

The authors have performed taxonomic analysis of the metagenomes from polluted and non-polluted episodes of fog events at Mount Tai. The study has described the key differences in the bacterial composition and associated it with the environmental factors. The results are interesting, however, several key points need to be addresses:

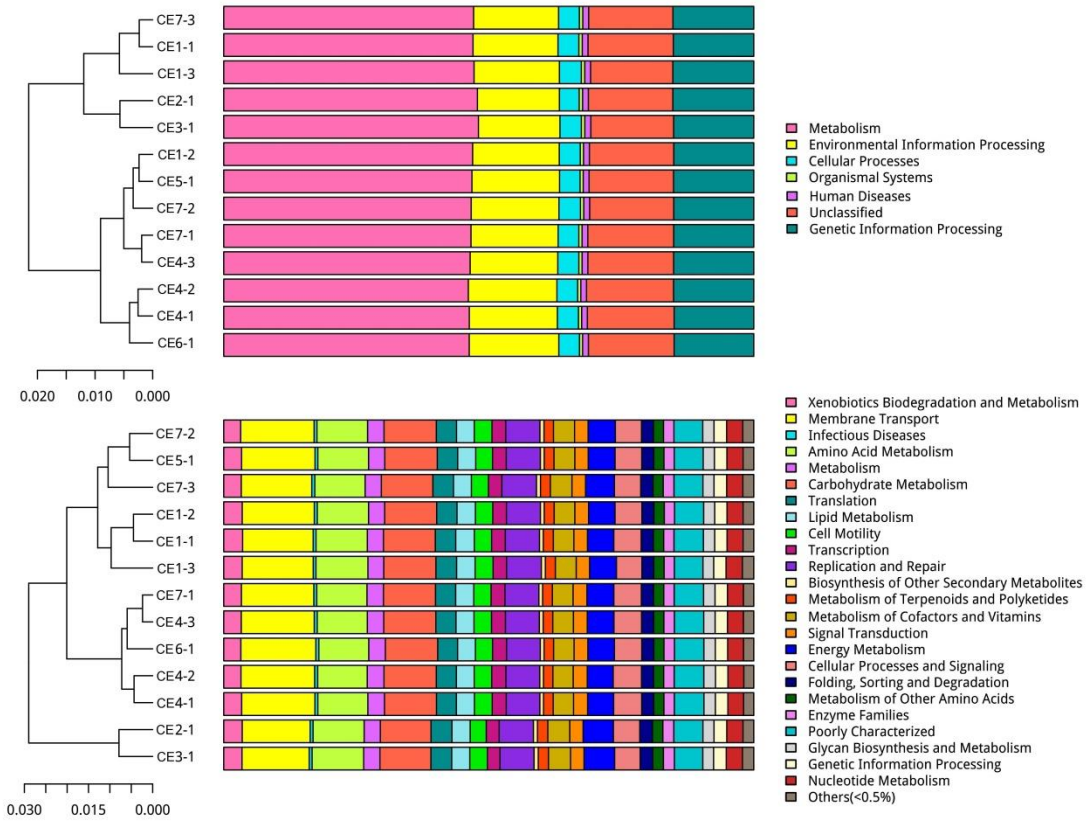
1. The authors performed a 16S analysis of the microbial community which is an excellent and reliable approach to classify the composition. However, the functionality of microbes cannot be assumed on this base. It would be other approaches (e.g. metagenome-assembly based analysis). Therefore, the authors should modify the text accordingly i.e. by not assuming the functional diversity of fog microbes. Please refer to Jiang et. al, Nat Protoc, 2015 for general methodology.

Response of the authors: thank you for your comments.
First, yes, the Miseq 16S rRNA gene sequencing was a powerful tool in microbial diversity investigation, which provides a comprehensive understanding of community composition. The metagenome-assembly based analysis have been widely used in microbial community functional analysis, e.g., Cao et al. (2014) described the microbial communities in $PM_{2.5}$ and PM_{10} using metagenomics during a serious smog event (Cao et al., 2014), Be et al examined aerosolized microorganisms in urban airborne microbes and revealed the metagenomic complexity of urban aerosols and the potential of genomic analytical techniques for biosurveillance and monitoring of threats to public health (Be et al., 2015). We also studied the suggested reference Jiang et. al for general methodology (Jiang et al., 2015). However, due to the

complexity of cloud water collection, the amount for each cloud episode ranged from 40 to 200 mL based on the duration and characteristics of the clouds. The sampled volume was inadequate for metagenomic analysis.

Second, community functions are based on community composition, bacterial taxa. For specific functional bacteria, e.g. Rhizobia (*Phyllobacterium myrsinacearum*) are involved in Biological nitrogen fixation, and favorable for plant growth, Methanotrophic Bacteria (*Methylobacterium aquaticum*, *Methylobacterium adhaesivum*) are related to methane oxidation and carbon recycling. We have added the references attached to the bacterial species in Table 3.

Third, to acquire more accurate community function, we performed PICRUST function prediction in the revised manuscript in section 3.1 (page 11). Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUST) can be used to predict the metabolic function spectrum of corresponding bacteria and archaea based on the 16S rRNA gene sequence (Langille et al., 2013). PICRUST has been used in bacterial diversity and function analysis (Corrigan et al., 2015; Wu et al., 2016) and it was applied in community function prediction in the present study. The predicted function including metabolism and human disease are essential part of atmospheric microbial community function, which are likely attributed to the bacterial gene content from the identified species in Table 3.



Predictive functional profiling of microbial communities, using 16s rRNA gene sequences

2. Quality of English needs to be revised. Grammatical mistakes and use of wrong terminologies is seen throughout the manuscript. For instance, grammatical mistake

in introduction (Line 5), Page 9 line 33 (“: : : server infections”) and Page 10 line 5 (“Previous studies has shown : : :”). The author names also have discrepancy in main text and supplement.

Response of the authors: We have polished the manuscript with a professional assistance in writing. The mistakes have been corrected.

3. Few major statements are not supported by references e.g. the authors mentioned that the recruitment of bacteria from various sources to air is harmful to humans upon deposition back to land via fog.

Response of the authors: Numerous studies are focused on the potential pathogens identified in atmospheric particulate matter ($PM_{2.5}$ and PM_{10}) (Cao et al., 2014; Creamean et al., 2013), rain water (Kaushik & Balasubramanian, 2012; Simmons et al., 2001), and indicated that health risk-related bacteria in atmospheric samples should be concerned. For cloud water, studies of health risks to individuals are typically focused on the chemical characteristic, e.g. the low pH (Hackney et al., 1989), PAH (Ehrenhauser et al., 2012), etc. Limited literatures discussed the microorganism in cloud water suggested the potential pathogens in cloud water (Vaïtilingom et al., 2012), they find potential plant pathogens such as *Pseudomonas syringae* and *Xanthomonas campestris*. In their study, they suggest that the wet deposition play a major role in the dispersion of microorganisms, which appear as an extension of the phyllosphere and carry living species of plant pathogens that could then infect new hosts through precipitation.
In the present study, potential human pathogens were identified in the cloud water samples. However, the detailed health risks should be prudently assessed. Further study depending on the culture-dependent method and biochemical experiments will perform to check the pathogenicity. We have revised the relevant discussion and added reference in page 12 (from line 25 to line 33), page 14 (from line 27 to line 35) and page 15 (from line 1 to line 14).

4. Figure captions need to a bit more descriptive/proof read.

Response of the authors: More detailed descriptions have been added to the figure captions.
Figure 1 Bacterial community variation for fog episodes at the phylum (A) and class (B) level. FE refers to the cloud episodes. Bar graphs for each sample represent the percentage of taxa assigned to each phylum with 80% bootstrap confidence. Taxonomic summary of the most abundant taxa (more than 1%) across all cloud samples are indicated in the figure.
Figure 2 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. Spearman's correlation coefficients were estimated for each pairwise comparison of

genus counts and KEGG pathway counts. Selected KEGG pathways relating to metabolism and disease infection and predominant genera are included in the heatmap. Red color refers to the positive correlation, and green indicates a negative correlation. Correlation is significant at *P < 0.05, **P < 0.01.

Figure 3 Schematic representation of bioaerosols life cycle and potential influence on atmosphere, ecosystem and human health, modified from Poeschl (Poeschl, 2006). The predominant bacteria species with potential functions are indicated in the figure. Bioaerosols emissions and resuspension from various terrestrial environments, e.g., soil, water, plants, animals or human beings, may include pathogenic or functional species. These bacteria can be attached to particles or incorporated into water droplets of clouds/fog. Certain species can serve as biogenic nuclei for Cloud Condensation Nuclei (CCN) and Ice Nuclei (IN), which induce rain formation, precipitation, and wet deposition of gases and particles. For the potential pathogens and functional bacteria, during cloud process, they can be deposited back to land via deposition and possibly induce infections to human health and impose effect on the diversity and function of aquatic and terrestrial ecosystems.

Figure 4 Bacterial taxa significantly differentiated between the polluted and non-polluted cloud episodes identified by linear discriminant analysis coupled with effect size (LEfSe). The LDA effect sizes (left) were calculated using the default parameters. The taxonomic cladogram (right) with LDA values higher than 3.5 comparing all bacterial taxa and significantly discriminant taxon nodes are colored and branch areas are shaded according to the highest-ranked variety for that taxon. Taxa with significant difference in polluted and non-polluted cloud episodes are indicated in red and green circles, respectively. The bacterial taxa with nonsignificant differences are represented as yellow circles and the diameter of the circles are proportional to relative abundance. The abbreviation in the cladogram tree: a: g_Rhodococcus, b: f_Nocardiaceae, c: f_Flavobacteriaceae, d: o_Flavobacteriales, e: c_Flavobacteriia, f: p_Bacteroidetes, g: f_Deinococcaceae, h: o_Deinococcales, i: c_Deinococci, j: p_Deinococcus-Thermus, k: f_Bacillaceae, l: o_Clostridiales, m: c_Clostridia, n: p_Firmicutes, o: f_Oxalobacteraceae, p: o_Burkholderiales, q: c_Betaproteobacteria, r: f_Moraxellaceae, s: o_Pseudomonadales, t: f_Xanthomonadaceae, u: o_Xanthomonadales, v: c_Gammaproteobacteria, w: p_Proteobacteria.

Figure 5 Biplot of the environmental variables and genus-level community structure using a redundancy analysis (RDA) model, describing the variation in bacterial community explained by environmental variable. CE refers to cloud episodes. Polluted episodes are indicated in red circle, and non-polluted episodes are green squares. Species data are listed in Table S2. The selected environmental variables are significant (P < 0.05) using Monte Carlo permutation testing. Species are labeled with triangle, the closer of which indicates existence in similar environment. Environmental variables are showed by arrows; the relative length is positive correlation with the importance in influencing bacterial community structure. The angle between the arrow and the ordination axis suggest the variable response respect to the RDA gradient. The two axes explain 65.9% of the variability. For bacteria,

major ions in cloud water seem to be the most important environmental variable shaping the community structure.

Figure 6 Air mass transport pathways for the cloud episodes using the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model. 24-hour backward trajectories were calculated for air parcels arriving at the summit of Mt. Tai (36o18′ N, 117o13′ E, and 1534 m a.s.l). CE refer to cloud episodes. The polluted episodes are indicated in red lines, and green lines are non-polluted episodes.

Figure 7 Wind Rose Diagram to quantitative analysis of wind speed and wind direction during sampling time between polluted (A) and non-polluted cloud episodes (B). The frequency of winds is plotted by wind direction, the color bands show wind speed range. The direction of the longest spoke shows the wind direction with the greatest frequency. Figure 7 (C) indicates distribution of wind speed during sampling time and correlation with $PM_{2.5}$ concentration. As shown in the figure, $PM_{2.5}$ concentration was high under lower wind speed, whereas $PM_{2.5}$ was lower when wind speed was high.

Response to reviewer 2

The authors investigated the differences in bacterial community structures from fog water droplet samples collected from Mt. Tai in North Plain of China including those clear and polluted days in July and August of 2014. They performed sequence analysis of the samples, and also investigated the effects of environmental factors on the bacterial community structure. Overall, it is interesting to study the bacteria in the fog water samples, especially in higher altitude from a ground. The information developed is useful to understanding the microbial transport and possible roles in atmospheric pollutant transformation. The authors provided a number of different analyses of their results and derived some valuable information. Nonetheless, this reviewer does observe the following drawbacks that need the authors' attention:

1. From their work, it seems they only had one day with higher PM2.5 pollution level, i.e., exceeding 100 ug/mˆ3, and they had more samples from clear days with much lower PM2.5 levels. In their work, they compared them and further derived relevant information. I think the authors have to carefully make their conclusions regarding their limited set of data from a single polluted day. Probably, they can use the 24-hour backward trajectories to discuss more on them.

Response of the authors: According to our field observation data on the summit of Mt.Tai in the year of 2014 and 2015(unpublished), the 24 h average $PM_{2.5}$ mass concentration was basically less than 100 μg/m3, sometimes with relative lower concentration less than 10 $\mu g/m^3$. The possible reason was that the Mt. Tai was the highest mountain in North China Plain (1534 m a.s.l), which was typically as the background of atmospheric quality. The $PM_{2.5}$ was relatively low than other regions in the North China Plain. Similar results were obtained by other studies and suggest the relation of $PM_{2.5}$ and attitude. Gehrig and Buchmann studied the seasonal

variations and spatial distribution of ambient PM_{10} and $PM_{2.5}$ concentrations. In comparison to other study area (different attitude), the lowest $PM_{2.5}$ concentrations were observed at the elevated site Chaumont (1140 m a.s.l.) (Gehrig & Buchmann, 2003). Similarly, Fan et al studied the vertical distribution of $PM_{2.5}$ concentration in fog and haze days in Beijing and suggest that $PM_{2.5}$ concentrations decreased with the increase of altitude (Fan et al., 2009).

The 24 h average $PM_{2.5}$ mass concentration according to our field observation data

2014	PM2.5 (μg/m^3)	2014	PM2.5 (μg/m^3)	2015	PM2.5 (μg/m^3)	2015	PM2.5 (μg/m^3)
2014/7/23	53.9	2014/8/9	30.1	2015/7/6	62.3	2015/7/23	34.8
2014/7/24	49.9	2014/8/10	67.5	2015/7/7	64.8	2015/7/24	34.2
2014/7/25	12.2	2014/8/11	51.1	2015/7/8	112	2015/7/25	36
2014/7/26	44.5	2014/8/12	45.5	2015/7/9	71.6	2015/7/26	33.7
2014/7/27	66.9	2014/8/13	47.4	2015/7/10	61.2	2015/7/27	65.6
2014/7/28	97.5	2014/8/14	49.3	2015/7/11	80.6	2015/7/28	49
2014/7/29	73.4	2014/8/15	47.6	2015/7/12	31.4	2015/7/29	7.9
2014/7/30	23.4	2014/8/16	66.2	2015/7/13	57	2015/7/30	12.6
2014/7/31	56.2	2014/8/17	69.5	2015/7/14	53.5	2015/7/31	14.9
2014/8/1	17.7	2014/8/18	53.7	2015/7/15	52.1	2015/8/1	17.2
2014/8/2	42.5	2014/8/19	64.9	2015/7/16	54.5	2015/8/2	35
2014/8/3	45.5	2014/8/20	62.1	2015/7/17	62.9	2015/8/3	6.1
2014/8/4	86.5	2014/8/21	71.3	2015/7/18	51.9	2015/8/4	5.9
2014/8/5	54.8	2014/8/22	54.2	2015/7/19	44.4	2015/8/5	9.9
2014/8/6	18.2	2014/8/23	48.8	2015/7/20	21.4	2015/8/6	19.3
2014/8/7	44.1	2014/8/23	42.8	2015/7/21	43.1	2015/8/7	7.8
2014/8/8	40.3			2015/7/22	55.8	2015/8/8	26.2

In addition, the listed $PM_{2.5}$ concentration in Table 1 was the average value during a cloud process, not the 24 h average concentration. The 24 h $PM_{2.5}$ concentration in cloud days was lower than non-cloud days which possible due to the wet deposition. During cloud episodes, $PM_{2.5}$ concentration varied with cloud process. The mass concentration was high in the initiation of cloud episode, with the development and dissipation of cloud, the concentration steadily reduced due to the reduced input (nighttime) and wet deposition.

In the present study, the polluted cloud episodes were firstly defined according to the 24 h concentration of WHO air quality guideline (25 μg/m^3) and the standard was applied by Australia, New Zealand and European Union. In the revised manuscript, we checked the major ions in cloud water and reclassify the cloud episodes according to the concentration of major ions in cloud water. Cloud episodes under high concentration of ions in cloud water and high atmospheric $PM_{2.5}$ concentration were defined as polluted. The cluster and PCA analysis also confirmed the classification.

In the section of 3.4, we have discussed the influence of air mass and meteorological conditions on $PM_{2.5}$. The Sampling site was 1534 m a.s.l, air pollution was typically effected by air mass over long term transport than local emissions.We use the 24-hour backward trajectories to track the air mass and combined the wind direction and wind

speed to deeply discuss the possible driven factors. The main points obtained was that air mass from the contaminated area through long term transport with lower wind speeds, largely reduce the diffusion rate of pollutants and thus lead to the sustained high $PM_{2.5}$ during polluted cloud episodes.

2. It seems they did not clearly define what level of PM2.5 for which a day can be classified as a polluted day in their method section. Also they should clearly define what those symbols such as "FE" stand for? although I guess it should be "Fog Episode", but they should appear in all figure captions so that readers can easily understand the figures. They should describe that the characteristics of each Fog Episode are shown in relevant Tables in each Figure.

Response of the authors: thank you to your suggestion, we have replaced “fog episodes” as “cloud episodes” and described cloud episode in relevant Tables and Figures.
We have clearly defined the polluted cloud episodes in page 5 (from line 21 to line 29), page 9 (from line 14 to line 26) and indicated the abbreviation in all the table and figures. The polluted cloud episodes were firstly defined according to the 24 h concentration of WHO air quality guideline (25 $\mu g/m^3$) and the standard was applied by Australia, New Zealand and European Union. During a cloud episode, the average $PM_{2.5}$ concentration higher than 25 $\mu g/m^3$ was classified as polluted. WHO proposes $PM_{2.5}$ less than 10 ug/m^3 is safe. Elevated $PM_{2.5}$ concentration will highly increase health risks. The high pollutant and pathogens are detrimental to individuals (Fang et al., 2013). $PM_{2.5}$ concentrations were compared to the 24 h World Health Organization limit of 25 $\mu g/m^3$.
In the revised manuscript, we checked the major ions in cloud water and reclassify the cloud episodes according to the concentration of major ions in cloud water. Cloud episodes under high concentration of ions in cloud water and high atmospheric $PM_{2.5}$ concentration were defined as polluted. The cluster and PCA analysis also confirmed the classification.

3. be aware that they only performed genus level sequence and they cannot derive any particular bacterial species, especially when they discuss about pathogens. For certain genera, not all of their species are pathogens or opportunistic pathogens.

Response of the authors: we have revised the discussion about potential pathogens in page 12 (from line 25 to line 33), page 14 (from line 27 to line 35) and page 15 (from line 1 to line 14). Yes, the Miseq sequencing can identify bacterial taxa mostly at the genus level. In the present study, the V3-V4 hypervariable region of the bacterial 16S rRNA gene was amplified. Single OTUs were removed and taxonomy was assigned to each OTU using the Ribosomal Database Project (RDP) classifier in QIIME, with a minimum confidence cutoff of 80% against the Silva reference database (silva 119,http://www.arb-silva.de/) to the genus level. Finally, we focused on those bacterial genera that included species known or suspected to be opportunistic

pathogens. To this aim, we performed a systematic literature review to identify potential pathogenic bacteria in water habitats (Bibby et al., 2010; Guo & Zhang, 2012; Luo & Angelidaki, 2014).
Previous studies have discussed the bacterial pathogens based on NGS sequencing (454 pyrosequencing, Miseq, Ion Torrent PGM). Razzauti et al conducted a comparison between transcriptome sequencing and 16S metagenomics for detection of bacterial pathogens in Wildlife (Razzauti et al., 2015) and suggest that 16S approach was able to determine bacterial diversity in each individual. They also indicated that NGS techniques (454-pyrosequencing and MiSeq) are very affordable candidates and could become routine approaches in future large-scale epidemiological studies. Luo and Angelidaki studied the bacterial communities and bacterial pathogens with the high sequencing depth by Ion Torrent PGM (16S rRNA gene sequencing), they suggest the Ion Torrent PGM is also possible to detect the potential bacterial pathogens in biogas reactors. To identify potential pathogens, they use the reference bacterial pathogen database and identified the potential bacterial pathogens at the species level (Luo & Angelidaki, 2014).

4. I did not see any concentration levels for the total bacteria in their fog water droplet samples? Did they perform qPCR for total bacteria for their samples?

Response of the authors: Due to the complexity of cloud water collection, the amount for each cloud episode ranged from 40 to 200 mL based on the cloud duration and characteristics. For the majority episodes, e.g. CE1-3, CE2-1, CE4-1, CE4-2, the remained volume was inadequate for other analysis after Miseq sequencing.
The collected cloud water samples were processed by genomic DNA extracting, PCR amplification, Miseq sequencing and qPCR. In DNA extraction, some samples DNA cannot be successfully extracted and require repeated extraction, thus consume more sample volume. We have performed qPCR for total bacteria after Miseq. However, after miseq, no remaining sample DNA for the further analysis for certain samples. QPCR was just performed for the samples with sufficient DNA and bacterial concentration are listed in the following table. Therefore, we did not discuss the total bacterial concentration in the manuscript.

Bacterial concentration for different cloud episodes

Sample	Collected volume (mL)	Bacterial concentration (cells/mL)
CE1-1	90	8.9×10^{4}
CE1-2	80	1.3×10^{5}
CE1-3	55	Not detected
CE2-1	75	Not detected
CE3-1	100	Not detected
CE4-1	65	Not detected
CE4-2	40	Not detected
CE4-3	40	Not detected
CE5-1	50	Not detected

CE6-1	60	Not detected
CE7-1	210	1.5×10^5
CE7-2	200	5.8×10^4
CE7-3	120	1.6×10^5

5. It would be great if they can provide data for fungal spores. I guess there will be some fungal spores in the fog water droplets.

Response of the authors: We agree that the investigations of fungal diversity in fog/cloud water are very important areas for future work. Your suggestions are very helpful for our further study. However, the analysis of fungal spores requires substantial amounts of additional work, including the resequencing and culture-dependent experiments. The remaining parts of the samples are unable to support the above experiments. We therefore decide not to include these in our current manuscript and leave for future work. The next studies on microbial community will consider the fungal diversity in fog/cloud water and other aerosol samples.

6. For certain bacteria, when they are stored at 4 degree C, they can still grow. How long did it elapse between the collection and their actual analysis?

Response of the authors: thank you for your comments. We have modified the description and clearly described the storage conditions of the sample in page 5(line 32) and page 6 (from line 1 to line 2). Basal analysis of water typically included chemical and biological two parts. For chemical analysis, part of samples were stored in pre-baked glass bottles, immediately preserved with hydrochloric acid (HCl, pH <2.0), stored at 4 ℃ in ice box during transit, and analyzed upon arrival at the Laboratory. Samples for microbial diversity analysis were not preserved with hydrochloric acid and stored with dry ice in transit, and frozen at -80 ℃ until further analysis.

Response to reviewer 3

Min et al examine bacteria present in cloud water samples collected at Mt. Tai, China. They use a variety of techniques to examine the community composition of bacteria in the samples and attempt to assess differences as a function of a variety of environmental parameters, especially fine particle concentration levels. While the dataset is interesting and the work novel, I have numerous concerns about the work and its presentation.

Major comments:

1. The authors never make it very clear why they are examining bacteria in clouds (they are looking at clouds, not fog – see below). They talk about the importance of interaction with fog, but don't clarify why such interactions are important. They speak about deposition in clouds, but why is this really important if such bacteria would be

deposited anyway by wet or dry processes? Bacteria in cloud drops get there through scavenging of aerosol particles that are either themselves bacteria or have bacteria attached. Why, then, is it important to look at bacteria in cloud water? Why not look at them directly in PM2.5 samples? This would allow a much larger dataset to be examined, which would greatly help statistical analyses of relationships with environmental variables. For example, if one is interested in examining changes in bacterial populations with PM2.5 levels, it would be much more straightforward to look at bacteria directly in PM2.5.

Response of the authors: According to your suggestion, we have revised the manuscript in the introduction (page 3, line 29-32; page 4, line 1-15) and defined the samples as clouds.

In recent years, Northern China experienced serious air pollution. Mt. Tai (36°18′N, 117°13′E, and 1534 m a.s.l), locates on the summit of North China Plain, is frequently attacked by cloud events (Guo et al., 2012; Liu et al., 2012; Wang et al., 2011). In contaminated area, cloud typically contains numerous pollutants, e.g., sulfate and nitrate ions, organic carbon compounds, and bioaerosols. During cloud process, atmospheric bacteria attached to particles or incorporated in cloud droplets will be deposited back to land via wet deposition, which may induce health risks through microbial pathogens dispersion and potential effect on the diversity and function of aquatic or terrestrial ecosystems. Previous literature discussed the microorganism in cloud water suggested the potential pathogens in cloud water (Vaïtilingom et al., 2012). They find potential plant pathogens such as *Pseudomonas syringae* and *Xanthomonas campestris*. They also suggest that the wet deposition play a major role in the dispersion of microorganisms, which could then infect new hosts through precipitation. Therefore, to evaluate the potential ecological functional bacteria in cloud water have been an urgent issue, especially for the polluted cloud episodes.

In cloud, bacteria can act as efficient cloud condensation nuclei or ice nuclei and associate with biogeochemical cycling (nitrogen/carbon cycling), which have been demonstrated by numerous studies (Amato et al., 2015; Hill et al., 2007; Vaïtilingom et al., 2013). Similarity, previous literatures have studied the bacterial community in rain or snow (Cho & Jang, 2014; Mortazavi et al., 2015). They also focus on the bacteria associated with CNN/IN, potential pathogens and biochemical reactions. In cloud, bacterial activities were essential to atmospheric hydrogenic cycle. To understand the bacterial community structure was the first step for further study on their activities in cloud. Therefore, we should first investigate bacterial community in cloud.

In the revised manuscript, we discussed the relationship between $PM_{2.5}$, major ions and bacteria. Major ions in cloud have vital role in bacterial community structure variation. In our opinion, major ions have direct influence on bacterial community, whereas $PM_{2.5}$ had an indirect impact on bacterial community structure.

2. One might be interested in examining how cloud processing affects bacteria. For example, do they differentially scavenge and deposit bacteria from a certain subset of

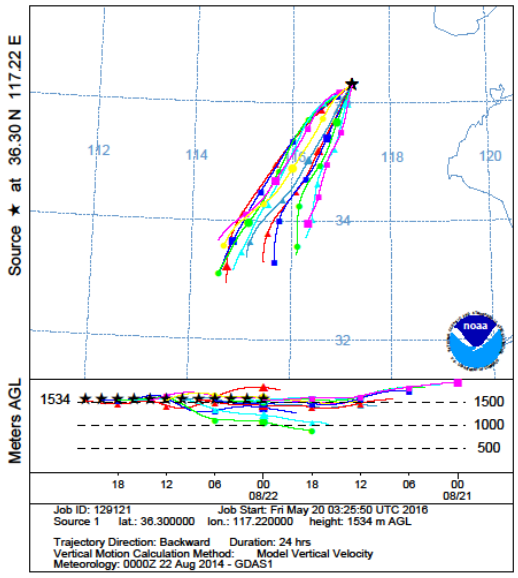
aerosol particles? Do the bacteria reproduce in clouds as suggested by Fuzzi? Does interaction with fogs alter the viability of bacteria in some way. The authors do not examine any such questions that would be very relevant to bacteria in fog.

Response of the authors: Your suggestions are very helpful for our further research. All the suggested researches are based on the in situ analysis during sampling and culture-dependent experiments in lab. We will perform the culture-dependent methods in laboratory to check the bacterial activity in further study, such as isolation of viable bacteria under low temperature and drop-freezing assays.
In the present study, we tell the readers which bacteria were in cloud water and whether there was different between polluted and non-polluted cloud by investigation of community structure. The basic study will provide fundamental acquaintance for the further research on bacterial activity in cloud.

3. I have many concerns about the way in which the authors assess differences in bacteria in fog between polluted and nonpolluted conditions. Chief among these is their classification of clean and polluted fog episodes. If one examines the back-trajectories, one finds very similar transport patterns in some cases for polluted and non-polluted cases. Furthermore, one can even find sequential samples within a single fog episode that are classified as clean and as polluted. Episode 7 is a good example, where sample 1 is classified as polluted, sample 2 is clean, and sample 3 again polluted. As shown in Figure 7, these samples all have essentially the same transport pattern. It is completely unreasonable to make such a separation based on $PM_{2.5}$ concentration, especially since the measured $PM_{2.5}$ in fog does not represent the actual fine particle load upon which the cloud formed since many particles are scavenged in fog and not, therefore, measured by the $PM_{2.5}$ monitor inside a cloud.

Response of the authors: we have redefined the polluted and non-polluted cloud episodes in the revised manuscript. Polluted cloud episodes were defined based on concentration of $PM_{2.5}$ and major ions in cloud water not just based on $PM_{2.5}$ concentration.
In continues cloud episodes, back-trajectories may be similar transport patterns. However, $PM_{2.5}$ concentration varied during cloud process. We measured the atmospheric $PM_{2.5}$ concentration and the variation of $PM_{2.5}$ was caused by wet deposition during cloud process. In the cloud formation stage, $PM_{2.5}$ concentration was high. With the development, the soluble components in $PM_{2.5}$ were scavenged in cloud. Atmospheric $PM_{2.5}$ gradually decreased. However, new pollutant from local regional emission will cause the elevated $PM_{2.5}$ (see cloud episode 7). For instance, in cloud episode 7, $PM_{2.5}$ concentration increased in the dissipation stage from 6:00 to 9:00 in the morning. The busy traffic vehicle and the industrial and agricultural activities resulted in the emissions of new pollutants into atmosphere. Therefore, an increased $PM_{2.5}$ concentration was observed.
In the revised manuscript, we have defined the polluted and non-polluted cloud episodes based on the major ions and $PM_{2.5}$ concentration. Cloud episodes under high

concentration of ions in cloud water and high atmospheric $PM_{2.5}$ concentration were defined as polluted.



	CE7-1	CE7-2	CE7-3
Time	2:30-4:38	4:38-6:21	6:21-9:20
$PM_{2.5}$($\mu g \cdot m^{-3}$)	30.45	23.39	41.60
EC ($\mu S \cdot cm^{-1}$)	356.20	207.50	187.60
Cl^-(μeq^{-1})	182.61	47.50	83.92
NO_3^-(μeq^{-1})	581.20	281.20	330.74
SO_4^{2-}(μeq^{-1})	2391.69	1045.64	1633.34
Na^+(μeq^{-1})	70.42	31.33	38. 40
NH_4^+(μeq^{-1})	2599.56	1721.15	2412.55
K^+(μeq^{-1})	85.77	42.20	61.15
Mg^{2+}(μeq^{-1})	156.02	72.57	68.03
Ca^{2+}(μeq^{-1})	963.70	417.79	387.14

4. Further issues regarding the author's classification of fog samples are apparent in the various attempts to statistically compare bacterial composition across fog samples. Looking at fog episode 7 again, as one example, one finds samples 1, 2, and 3 end up in very different clusters in Fig. 2. Likewise sequential "clean samples" 1-2 and 1-3 cluster very differently. These observations suggest to me that the author's approach may not be getting at real differences driving bacterial populations.

Response of the authors: The classification of fog samples were first based on the 24 h concentration of WHO air quality guideline (25 $\mu g/m^3$) and this standard has been applied by Australia, New Zealand and European Union. During a fog episode, the average $PM_{2.5}$ concentration higher than 25 $\mu g/m^3$ was classified as polluted.

In the revised manuscript, we checked the major ions in cloud water and reclassify the cloud episodes according to the concentration of major ions in cloud water.

For cloud episode 7, as mentioned in question 3, $PM_{2.5}$ and major ions dymaic with cloud process. In the cloud formation stage, $PM_{2.5}$ concentration was high. With the development of cloud, the soluble components in $PM_{2.5}$ were scavenged in cloud. Atmospheric $PM_{2.5}$ gradually decreased. However, new pollutant from local regional emission will cause the elevated $PM_{2.5}$ (see cloud episode 7). The varied $PM_{2.5}$ and major ions concentration were responsible for the different distribution. The concentration of major ions in the samples of CE7-1 and CE7-3 were higher than in CE7-2. Sample CE7-2 was cleaned with lower $PM_{2.5}$ and ions concentration.

Although the $PM_{2.5}$ concentration for sample CE1-2 was low, a relative high major ions concentration was detected in cloud water. Therefore, we categorized the sample CE1-2 as polluted sample. The cluster and PCA analysis also confirmed the classification (Figure S4), CE1-2 was closely to other polluted samples.

5. The manuscript lacks adequate description of sampling methodology. One important issue when measuring cloud composition is how the cloud collector is cleaned. This is particularly true for biological sample characterization as attempted here. How was the cloud collector cleaned? Was it sterilized? Was it cleaned just prior to each cloud event? Was the collector kept closed prior to cloud interception to ensure it did not become contaminated? Were cloud collector blanks taken? What bacteria were found in blanks? How do these relate to bacteria observed in samples? Without such information one cannot trust the measured bacteria to have come only from the cloud and not from the sampler.

Response of the authors: we have added the description of sampling methodology in page 5(from line 8 to line 20). The cloud collector was cleaned prior to each cloud event and kept closed prior to cloud interception to ensure not to be contaminated. The collector was activated by a sensor only when cloud formed in the ambient air. The cloud water was aspirated through a Teflon duct at a rate of 24.5 m^3 min^{-1} by a fan situated at the rear of Caltech Active Strand Cloud water Collector (CASCC) (See the attached figure). Cloud water collected from the Teflon strands, through Teflon tube and down into Teflon bottles. The theoretical 50% cut-off size was equivalently drop diameter of 3.5 μm.

To avoid the artificial and instrumental contamination, the Teflon tube and the polyethylene bottles were first pretreated with anhydrous ethanol and then washed 3 times using the sterilized ultrapure water. Before sampling, the cloud collector was washed with the sterilized deionized distilled water filtered through the with 0.22 μm membrane. Then spray the sterilized dd-H_2O into the collector using the sprayer and the collected water sample was as the blank. The parallel aliquots of sterile dd-H_2O were run through an identical DNA extraction procedure to check for sample contamination. These DNA extraction "blanks" were PCR amplified alongside the DNA samples extracted from the cloud water samples. Genomic DNA cannot be extracted from the blank samples. Therefore, we considered the cloud collector was strictly sterilized and cleaned before sampling and cannot contaminate the collected cloud water.

6. The manuscript is not well written. Grammar and syntax are very poor. At many points the authors' use of English language makes it difficult for the reader to even understand their meaning. Looking closely just at the abstract I counted more than 20 corrections needed to the text and several instances where the authors' meaning was unclear. I did look at some of the manuscript changes recently posted by the authors in response to other reviewer comments and found some improvements to the manuscript text but still observed many problems with the language.

Response of the authors: We have polished the manuscript.

Minor comments:

A. The cloud collector is not properly described. A CASCC2 has a flow rate below 5 m3/min. The 24 m3/min flow rate specified corresponds to a CASCC collector. See

collector descirptions and flow rates in Demoz et al. (1996) On the Caltech Active Strand Cloudwater Collectors. Atmos. Res., 41, 47-62.
Response of the authors: The CASCC collector was used during cloud water collecction, we have corrected.

B. More information needs to be provided about the trajectory calculations. What heights were used as trajectory endpoints?

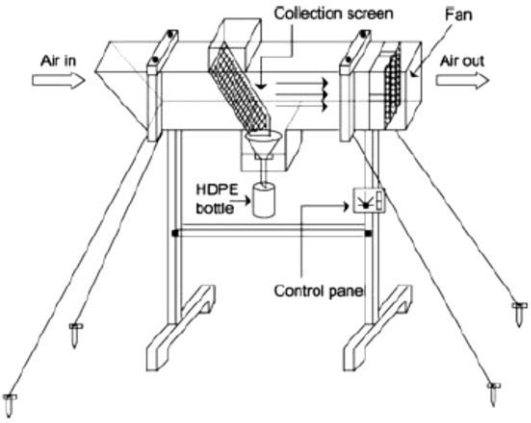
Response of the authors: We have indicated the location and heights in figure. The height for the backward trajectories was 1534 m (the height of sampling site on the summit of Mt. Tai).

C. More information should be given about sample handling. The biological samples should have been frozen, not refrigerated at 4 C. How much sample was collected? How much was used in the DNA workup?

Response of the authors: We have answered the question in the response to review 1. Basal analysis of water typically included chemical and biological two parts. For chemical analysis, part of samples were stored in pre-baked glass bottles, immediately preserved with hydrochloric acid (HCl, pH <2.0), stored at 4 ℃ in ice box during transit, and analyzed upon arrival at the Laboratory. Samples for microbial diversity analysis were not preserved with hydrochloric acid and stored with dry ice in transit, and frozen at -80 ℃ until further analysis.
The collected sample volume has been listed in the Table (response to reviewer 2, question 4). The collected volume for each sample was varied across the 13 cloud water samples and ranged from 40 to 200 mL. For each sample, at least 20 mL cloud water sample was used in DNA workup.

D. Some of the fog collection periods were quite long – up to 9 hrs. Was the fog continuously present during this entire period? If not, collected fog water could evaporate and aerosol particles could be captured on collector surfaces, contaminating the fog sample.

Response of the authors: The collection of cloud water was continuously during the entire period. As shown in the following figure, the polyethylene bottle was tight with the lid to avoid evaporation and external contamination. Cloud water in air was aspirated through the duct. To avoid the contamination from the aerosol particles and rainwater, a triangular roof was in the front of the duct.



E. It would be helpful to include additional information about the fog samples? At a minimum, the authors should include standard parameters such as cloud liquid water content during the sample, concentrations of major ions (which would provide greater insight into pollution levels), and cloud water total organic carbon.

Response of the authors: we have added the data of cloud liquid water content, the concentration of major ions and total organic carbon in cloud water in the Supplement Materials (Figure S1). We also analyzed the correlation between these environmental parameters and bacterial community in the manuscript in the discussion section about bacterial community and environmental factors.

F. The water samples collected atop Mt. Tai in summer are almost certainly associated with intercepted clouds. I suggest the authors not refer to these as fogs.

Response of the authors: we have revised the description and replaced “fog” as “cloud”.

Reference

Amato, P., Joly, M., Schaupp, C., Attard, E., Möhler, O., Morris, C.E., et al.: Survival and ice nucleation activity of bacteria as aerosols in a cloud simulation chamber, Atmos. Chem. Phys., 15, 6455-6465, 2015.

Be, N.A., Thissen, J.B., Fofanov, V.Y., Allen, J.E., Rojas, M., Golovko, G., et al.: Metagenomic analysis of the airborne environment in urban spaces, Microb. Ecol., 69, 346-355, 2015.

Bibby, K., Viau, E., Peccia, J.: Bibby K, Viau E, Peccia J.. Pyrosequencing of the 16S rRNA gene to reveal bacterial pathogen diversity in biosolids. Water Res 44: 4252-4260, Water. Res., 44, 4252-4260, 2010.

Cao, C., Jiang, W., Wang, B., Fang, J., Lang, J., Tian, G., et al.: Inhalable microorganisms in Beijing's $PM_{2.5}$ and PM_{10} pollutants during a severe smog event, Environ. Sci. Technol., 48, 1499-1507, 2014.

Cho, B.C., Jang, G.I.: Active and diverse rainwater bacteria collected at an inland site in spring and summer 2011, Atmos. Environ., 94, 409-416, 2014.

Corrigan, A., de Leeuw, M., Penaud-Frezet, S., Dimova, D., Murphy, R.A.: Phylogenetic and

functional alterations in bacterial community compositions in broiler ceca as a result of mannan oligosaccharide supplementation, Appl. Environ. Microbiol., 81, 3460-3470, 2015.

Creamean, J.M., Suski, K.J., Rosenfeld, D., Cazorla, A., Demott, P.J., Sullivan, R.C., et al.: Dust and biological aerosols from the Sahara and Asia influence precipitation in the western U.S, Science, 339, 1572-1578, 2013.

Ehrenhauser, F.S., Khadapkar, K., Wang, Y., Hutchings, J.W., Delhomme, O., Kommalapati, R.R., et al.: Processing of atmospheric polycyclic aromatic hydrocarbons by fog in an urban environment, J. Environ. Monitor., 14, 2566-2579, 2012.

Fan, W., Hu, B., Wang, Y., Wang, S., Sun, Y.: Measurements on the Vertical Distribution of $PM_{2.5}$ Concentration in Fog and Haze Days in Beijing City, Climatic & Environmental Research (Chinese) 14, 631-638, 2009.

Fang, W., Yang, Y., Xu, Z.: PM_{10} and $PM_{2.5}$ and health risk assessment for heavy metals in a typical factory for cathode ray tube television recycling, Environ. Sci. Technol., 47, 12469-12476, 2013.

Gehrig, R., Buchmann, B.: Characterising seasonal variations and spatial distribution of ambient PM_{10} and $PM_{2.5}$ concentrations based on long-term Swiss monitoring data, Atmos. Environ., 37, 2571-2580, 2003.

Guo, F., Zhang, T.: Profiling bulking and foaming bacteria in activated sludge by high throughput sequencing, Water. Res., 46, 2772-2782, 2012.

Hackney, J.D., Linn, W.S., Avol, E.L.: Acid fog: effects on respiratory function and symptoms in healthy and asthmatic volunteers, Environ. Health Perspect., 79, 159-162, 1989.

Hill, K.A., Shepson, P.B., Galbavy, E.S., Anastasio, C., Kourtev, P.S., Konopka, A., et al.: Processing of atmospheric nitrogen by clouds above a forest environment, J. Geophys. Res., 112, 2864-2864, 2007.

Jiang, W., Liang, P., Wang, B., Fang, J., Lang, J., Tian, G., et al.: Optimized DNA extraction and metagenomic sequencing of airborne microbial communities, Nat. Protoc., 10, 768-779, 2015.

Kaushik, R., Balasubramanian, R.: Assessment of bacterial pathogens in fresh rainwater and airborne particulate matter using Real-Time PCR, Atmos. Environ., 46, 131-139, 2012.

Langille, M.G., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., et al.: Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences, Nat. Biotechnol., 31, 814-821, 2013.

Luo, G., Angelidaki, I.: Analysis of bacterial communities and bacterial pathogens in a biogas plant by the combination of ethidium monoazide, PCR and Ion Torrent sequencing, Water. Res., 60, 156-163, 2014.

Mortazavi, R., Attiya, S., Ariya, P.A.: Arctic microbial and next-generation sequencing approach for bacteria in snow and frost flowers: selected identification, abundance and freezing nucleation, Atmos. Chem. Phys., 15, 6183-6204, 2015.

Razzauti, M., Galan, M., Bernard, M., Maman, S., Klopp, C., Charbonnel, N., et al.: A Comparison between Transcriptome Sequencing and 16S Metagenomics for Detection of Bacterial Pathogens in Wildlife, Plos. Neglect. Trop. D., 9, 876-887, 2015.

Simmons, G., Hope, V., Lewis, G., Whitmore, J., Gao, W.: Contamination of potable roof-collected rainwater in Auckland, New Zealand, Water. Res., 35, 1518-1524, 2001.

Va ïilingom, M., Attard, E., Gaiani, N., Sancelme, M., Deguillaume, L., Flossmann, A.I., et al.: Long-term features of cloud microbiology at the puy de Dôme (France), Atmos. Environ., 56, 88-100, 2012.

Va ïilingom, M., Deguillaume, L., Vinatier, V., Sancelme, M., Amato, P., Chaumerliac, N., et al.: Potential impact of microbial activity on the oxidant capacity and organic carbon budget in clouds, Proc. Natl. Acad. Sci. USA., 110, 559-564, 2013.

Wu, J., Peters, B.A., Dominianni, C., Zhang, Y., Pei, Z., Yang, L., et al.: Cigarette smoking and the oral microbiome in a large study of American adults, ISME. J., 10, 2435-2446, 2016.

The revised manuscript

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

Min Wei [1], Caihong Xu[1], Jianmin Chen[1,2,*], Chao Zhu[1], Jiarong Li[1], Ganglin Lv [1]

[1] Environment Research Institute, School of Environmental Science and Engineering, Shandong University, Ji'nan 250100, China

[2] Shanghai Key Laboratory of Atmospheric Particle Pollution and Prevention (LAP), Fudan Tyndall Centre, Department of Environmental Science & Engineering, Fudan University, Shanghai 200433, China

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

Abstract:

Bacteria, widely distributed in atmospheric bioaerosols, are indispensable components in clouds and play an important role in atmospheric hydrological cycle. However, limited acknowledge is acquired about bacterial community structure and function, especially for the increasing air pollution events in North China Plain. Here we presented a comprehensive characterization of bacterial community composition, function, variation and environmental influence for the cloud water collected at Mt. Tai from 24 Jul to 23 Aug 2014. Using the Miseq 16S rRNA gene sequencing, the facts that cloud water harbored a highly diverse bacterial community and the predominant phyla of Proteobacteria, Bacteroidetes, Cyanobacteria and Firmicutes were investigated. The presence of bacterial taxa survived in low temperature, radiation and poor nutrients conditions were encountered in cloud water, suggesting well adaption to extreme environment. Bacterial gene functions predicted from 16S rRNA gene using the Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) suggested the pathways relating to metabolism and

disease infections are significantly correlated to the predominant genera. The abundant genera *Acinetobacter*, *Stenotrophomonas*, *Pseudomonas*, and *Empedobacter* originated from a wide range of habitat included cloud condensation nuclei and ice nuclei active species, opportunistic pathogenic and functional species, demonstrating the bacterial ecological and healthy importance in cloud water should be concerned. Clustering analysis including hierarchical cluster (Hcluster) and principal coordinate analysis (PCoA) indicated a significant disparity between polluted and non-polluted samples. Linear discriminant analysis effect size (LEfSe) demonstrated that the polluted cloud samples were enriched with potential pathogens. The non-polluted samples had more diverse ecological function groups. Community structure discrepant performed by redundancy analysis (RDA) indicated that major ions in cloud water and $PM_{2.5}$ have negative impact on bacteria, playing vital role in shaping microbial community structure. Major ions may provide available nutrition for bacteria and have direct influence on bacterial community, whereas $PM_{2.5}$ in air had an indirect impact on bacterial community structure. During wet deposition, soluble particulate matter dissolved in water droplets and resulted in the elevated concentration in cloud water. $PM_{2.5}$ was possibly associated with different origins and pathways of air mass using source tracking by the backward trajectory and wind analysis, mainly related to the long-term transport. This work furthered our understanding of bacterial ecological characteristics in the atmospheric aqueous phase, highlighted the potential influence of environmental variables on bacterial community over cloud process. It may provide fundamental acquaintance of bacterial community response in cloud water under increasing pollution stress.

Key words： cloud water, 16r RNA gene, function prediction, major ions, $PM_{2.5}$

1. Introduction

Cloud is the aerosol system composed of tiny droplets suspended in the atmosphere. In the atmosphere, pollutants attached to particles could be dissolved or incorporated into cloud droplets, which may induce complex impacts on environment security and human health. Over the past decades, studies on cloud water have mainly focused on the physicochemical properties (Aikawa et al., 2001; Boris et al., 2015; Fernández-González et al., 2014). Recently, with the in-depth understanding of the characteristics of cloud, studies on bioaerosols in clouds have been in the ascendant.

Studies have showed that living microorganisms, including bacteria, fungi and yeasts, are present in clouds (Burrows et al., 2009). As the first study on biological particles in fog/cloud water, Fuzzi et al (1997) suggest the bacterial replication in foggy days. Afterwards, with the development of detection techniques, microorganisms in fog/cloud water have been systematically studied (Amato et al., 2007c; Delort et al., 2010; Vaïtilingom et al., 2012). Combined with the field investigations and lab experiments, diverse bacterial communities are retrieved, and the bacterial metabolism active in cloud water are further demonstrated. In atmospheric aqueous phase, microorganisms can act as cloud condensation nuclei (CCN) and ice nuclei (IN), which have potential impact on cloud formation and precipitation processes (Amato et al., 2015; Bauer et al., 2003; Mortazavi et al., 2015). Moreover, microorganisms in cloud water are available to metabolize organic carbon compounds (degrading organic acids, formate, acetate, lactate, succinate) and associate with carbon and nitrogen recycling (Amato et al., 2007a; Hill et al., 2007; Vaïtilingom et al., 2010). They can also influence photochemical chemical reactions (Vaïtilingom et al., 2013) and participate in a series of complex and diverse biochemical metabolic activities.

Cloud occurrence is a complex process, in contaminated areas, cloud typically contains numerous pollutants, e.g., sulfate and nitrate ions, organic carbon compounds, and bacteria (Badarinath et al., 2007; Després et al., 2012; Fernández-González et al., 2014; Mohan & Payra, 2009). As an intensive agricultural and economic region in China, the North China Plain has been suffering serious air pollution in recent years, e.g., the severe fog and haze pollution in Beijing, Ji'nan in January 2013 (Huang et al., 2014; Wang et al., 2014). Mt. Tai (36°18′ N, 117°13′ E, and 1534 m a.s.l), the highest

mountain in North China Plain, is frequently attacked by cloud episodes (Guo et al., 2012a; Liu et al., 2012). Emissions and resuspension of bacteria by wind erosion or splashing water from various terrestrial environments into atmosphere recruit diverse airborne bacteria, which possibly involve opportunistic and functional bacteria. During cloud process, these bacteria attached to particles or incorporated in cloud droplets will be deposited back to land via dry or wet deposition. Accumulating literatures have indicated that bacteria in cloud/fog water droplets have potential effect on diversity and function of atmospheric and terrestrial ecosystems (Delort et al., 2010; Vaïlingom et al., 2013), even induce health risks through microbial pathogens dispersion (Vaïlingom et al., 2012). Previous literatures have studied the bacterial community in rain or snow (Cho & Jang, 2014; Mortazavi et al., 2015). They also focus on the bacteria associated with CNN/IN, potential pathogens and biochemical reactions. Therefore, to evaluate the potential ecological functional bacteria in cloud water have been an urgent issue, especially for the polluted cloud episodes.

It is noteworthy that atmospheric microorganisms are subject to a wide range of environmental condition including the meteorological factors and aerosols physiochemical composition (Womack et al., 2010). Community structure and function are closely related to the environmental characteristics in atmosphere and geomorphic characteristics (Dong et al., 2016; Gao et al., 2016). For instance, studies about inhalable bioaerosols in particulate matter suggest environmental parameters including temperature, relative humidity, PM_{10}, $PM_{2.5}$ and particle size have significant impact on the composition and dynamic of microbial communities (Adhikari et al., 2006; Bowers et al., 2013). However, due to the paucity of detailed and comprehensive studies of atmospheric bacterial composition, the understanding of the dynamic of bacterial community remains incomplete. During polluted cloud process, how bacterial community varied and which environmental factor play decisive role in shaping bacterial community structure are still scarcely studied.

In the present work, typical cloud episodes under polluted and non-polluted weather were collected on the summit of Mt. Tai in North China Plain. To understand the bacterial community structure and function, the Miseq 16S rRNA gene sequencing was performed, and PICRUSt predictive function was applied to examine the metabolic and ecological function. Analysis of similarities (ANOSIM) and linear discriminant analysis effect size (LEfSe) were executed to clarify the discrepant

bacterial taxa. Moreover, RDA analysis was applied to identify the pivotal environmental factor influencing bacterial community. Air mass back trajectory and wind analysis were conducted to definitude the most likely source and transmission paths of pollutants and bacteria.

2. Material and methods

2.1 Sample collection

Cloud samples were collected using the Caltech Active Strand Cloud water Collector (CASCC) on the summit of Mt. Tai (36°18′ N, 117°13′ E, and 1534 m a.s.l). The cloud collector was cleaned prior to each cloud event and kept closed prior to cloud interception to ensure not to be contaminated. The collector was activated by a sensor only when cloud formed in the ambient air. The cloud water was aspirated through a Teflon duct at a rate of 24.5 m^3 min^{-1} by a fan situated at the rear of collector. Cloud water collected from the Teflon strands, through Teflon tube and down into Teflon bottles. The theoretical 50% cut-off size was equivalently drop diameter of 3.5 μm.

To avoid artificial and instrumental contamination, the Teflon tube and the polyethylene bottles were first pretreated with anhydrous ethanol and then washed 3 times using the sterilized ultrapure water. Before sampling, the cloud collector was washed with the sterilized deionized distilled water filtered through the with 0.22 μm membrane. Then spray the sterilized dd-H_2O into the collector using the sprayer and the collected water sample was as the blank.

To distinguish the polluted and non-polluted cloud episodes, we firstly checked the air pollution condition according to the 24 h concentration of WHO air quality guideline (25 μg/m^3) and this standard has been applied by Australia, New Zealand and European Union. During a cloud episode, the average $PM_{2.5}$ concentration higher than 25 μg/m^3 was classified as air pollution. Further definition of cloud water was combined with the major ions in water droplets, which provide deep insight into pollution levels. Therefore, in the present study, cloud episodes under high atmospheric $PM_{2.5}$ concentration and high concentration of ions in cloud water were considered as polluted.

After adjustment, seven cloud episodes including thirteen samples were detected over the whole sampling period (from 24 July to 23 August 2014), including 11 polluted and 2 non-polluted cloud water samples (Figure S1). The samples for microbial

community investigation were stored with dry ice in transit and then frozen at -80^{o}C in laboratory until further analysis.

In cloud water, the pH and conductivity was detected with a Multi pH/COND/TEMP 6350 hand held Meter immediately after sampling. The major inorganic ions (Cl^{-}, NO_3^{-}, SO_4^{2-}, Na^{+}, K^{+}, Ca^{2+}, Mg^{2+}, and NH_4^{+}) in cloud water were quantified using the ion-chromatography system (Dionex ICS-90). Hourly data, e.g., meteorological parameters, and $PM_{2.5}$ were measured to evaluate the air quality during cloud episodes (Table 1). The meteorological parameters including atmospheric visibility, temperature, relative humidity, wind direction, wind speed were measured with an automatic meteorological station (PC-4, JZYG, China) *in situ*. The mass concentration of $PM_{2.5}$ was measured using a Model 5030 SHARP monitor (SHARP 5030, Thermo Fisher Scientific, Massachusetts, USA).

To determine the most likely source region for air mass of cloud episodes, the 24-h back trajectory analysis was performed using the Hybrid Single-Particle Lagrangian Integrated Trajectories (HYSPLIT) model (http://ready.arl.noaa.gov/HYSPLIT.php). Moreover, the wind rose diagram of study area (origin, version 9.0, Origin Lab Corporation, Northampton, MA) during cloud process were analyzed to clarify the predominant wind direction and wind speed.

2.2 DNA Extraction and PCR Amplification

Genomic DNA was extracted in triplicate with the FastDNA spin kit for soil (MP Biomedicals, Solon, OH, USA) according to the manufacturer's directions. The concentration of DNA was determined spectrophotometrically (Nano-Drop 2000, Thermo, Wilmington, Delaware, USA). To check sample contamination, DNA was extracted through an identical extraction procedure for the blank samples from the collected dd-H_2O. These blanks were PCR amplified together with the DNA samples extracted from the cloud water samples. For the blank, no obvious bands and target fragment was detected by examination of electrophoretic gel images.

The designed primer sets with the V3-V4 region of 16S rRNA gene (338F-806R) (Masoud et al., 2011), adapter and barcodes were selected in the illumina Miseq sequencing. For each sample, a 25-μL PCR mix was prepared containing 10 μL of 5x Buffer, 1μL of dNTP (10mM), 1 U Phusion High-Fidelity DNA polymerase, 20 ng of template DNA, 1 μL of each 10 μM modified primer, with double-distilled water until

25 μL. PCR was performed at 94 ℃ for 2 min; 25 cycles of 94 ℃ for 30 s, 56 ℃ for 30 s and 72 ℃ for 30 s; 72 ℃ for 5 min; and hold at 10 ℃.

The PCR products were separated by 2% agarose gel electrophoresis and purified with the nucleic acid purification Kit (AxyPrepDNA, Axygen, USA). Purified PCR products were quantified using a Qubit 3.0 fluorometer (Invitrogen, Carlsbad, CA) and then mixed to equal concentration. For each sample, 4 μL of 10 nM pooled DNA was denatured with 1 μL of 0.2 N NaOH at room temperature. Finally, the Illumina paired-end sequencing was performed on a MiSeq platform (Illumina, Inc., San Diego, CA). After sequencing, two FASTQ files (read 1 and read 2) for each sample were generated on the MiSeq reporter software automatically. Raw 16S rRNA gene sequences are available at the Sequence Read Archive (SRA) under accession number SRX1904235.

2.3 Illumina high-throughput sequencing and analyzing

Raw sequences were processed using the QIIME packages (Kuczynski et al., 2011). The pair-end reads were firstly merged with overlap length greater than 10 bp. Then, the adapter, barcodes and primers were removed from the merged sequences. Subsequently, the trimmed sequences with length shorter than 200 bp, quality score lower than 25, homologous longer than 8 bp, containing ambiguous characters were screened. Finally, chimeric sequences were distinguished using the Usearch61 algorithm and removed from the dataset. Optimized sequences were clustered into OTUs at the threshold of 97% similarity with the usearch61 algorithm. Single OTUs were removed and taxonomy was assigned to each OTU using the Ribosomal Database Project (RDP) classifier in QIIME, with a minimum confidence cutoff of 80% against the Silva reference database (silva 119, http://www.arb-silva.de/) to the genus level. Subsequently, we focused on the bacterial genera including species known or suspected to be opportunistic pathogen and performed a systematic literature review to identify potential pathogenic bacteria in water habitats (Bibby et al., 2010; Guo & Zhang, 2012b; Luo & Angelidaki, 2014).

To acquire bacterial community function, Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUST) was performed. The PICRUST can be used to predict the metabolic function pathway from corresponding bacteria and archaea and provide a community's functional capabilities based on the 16S rRNA

gene sequence (Langille et al., 2013). PICRUST has been used in bacterial diversity and function analysis (Corrigan et al., 2015; Wu et al., 2016). In the present, the phylogenetic and functional capacities for the bacteria in cloud water are of great interest to help understanding their roles in atmosphere, ecosystem and health. Bacterial community functional profiles were predicted from 16S rRNA gene using the PICRUSt program and annotated against with the Kyoto Encyclopedia of Gene and Genomes (KEGG) database. Spearman's correlation coefficients were calculated to link the pairwise comparison of KEGG pathway and genus. Selected KEGG pathways relating to metabolism and disease infection and predominant genera are included in the heatmap. Correlation is significant at P-value of less than 0.05 and 0.01.

Alpha diversity was assessed by examining the rarefaction curves, shannon-wiener curves and rank-abundance curves calculated with Mothur (v.1.34.0; http://www.mothur.org) (Schloss et al., 2009) and visualized in R project (v.3.1.3; https://www.r-project.org/). Community richness estimators including the observed OTUs, nonparametric Chao1, ACE, and community diversity estimators including Shannon and Simpson indexes were also calculated with Mothur. Moreover, the Good's coverage was used to evaluate the sequencing depth.

Principal component analysis (PCA) was performed to visualize the changes in bacterial community for the collected samples. The PCA plots were constructed depending on Bray-Curtis similarity index calculated with the abundance of OTUs using the Biodiversity package (Kindt & Coe, 2005) in R. The difference in OTU composition for samples collected in polluted and non-polluted cloud episodes was tested by the analysis of similarity (ANOSIM) (Clarke, 1993). ANOSIM was implemented with the VEGAN package in R. Linear discriminant analysis effect size (LEfSe, http://www.huttenhower.sph.harvard.edu/galaxy/) was applied to identify differentially abundant bacterial taxa associated with the polluted and non-polluted cloud episodes at genus or higher taxonomy levels (Segata et al., 2011). For all statistical tests, the P value less than 0.05 was considered significant.

2.4 Interaction between bacterial community structure and environmental variables

Correlation between bacterial community and environmental variables was first

performed using a detrended correspondence analysis (DCA) to estimate the length of the gradient. The resulting length (0.99) indicates a linear model was appropriate, hence RDA was subsequently performed. RDA was elaborated with the predominant bacteria data matrix and the environmental data matrix including $PM_{2.5}$ mass concentration, meteorological conditions, water pH, electric conductivity and major ions in cloud water (Anderson & Willis, 2008). Interset correlations were used to determine the most important environmental variables in determining the community structure. To explain the species data, cumulative fit per species as fraction of variance of species was analyzed. The importance for the ordination space and the species most associated with environmental factors were selected. Analysis was performed in Canonical Community Ordination for windows (Canoco, v 4.5).

3. Results and discussions

3.1 Definition of polluted and non-polluted cloud episodes

To distinguish the polluted and non-polluted cloud episodes, we firstly checked the air pollution based on the $PM_{2.5}$ concentration. In the present study, the average $PM_{2.5}$ concentration during a cloud episode higher than 25 $\mu g/m^3$ was classified as air pollution. We first defined cloud water samples according to the air pollution conditions. Cloud water collected under air pollution condition was considered as polluted cloud episodes. However, in PCoA and Hcluster analysis (Figure S3), sample CE1-2 considered as non-polluted was separated from the other non-polluted samples and closed to the polluted samples. Reclassification of cloud water samples were combined with the major ions in water droplets. By checking the major ions, we observed that although the $PM_{2.5}$ concentration for sample CE1-2 was low, a relative high major ions concentration was detected (Figure S1). Therefore, we categorized the sample CE1-2 as polluted sample. The cluster and PCA analysis also confirmed the classification (Figure S4).

Although the predominant bacteria are similar between polluted and non-polluted cloud episodes, significant disparity within bacterial taxa are also identified. ANOSIM analysis suggest that the OTUs from polluted samples were grouped into one large cluster, and separated from the non-polluted clusters (ANOSIM comparison, $R = 0.683$, $P = 0.012$, < 0.05). Cluster analysis including PCoA and Hcluster indicated a highly similar community composition in polluted samples, regardless of the cloud

episodes (Figure S3). Cluster analysis based on the relative abundance of genera showed similar clustering patterns (Figure S4), and the polluted samples also shared high similarity in their bacterial community structure.

3.2 Microbial community in cloud water

Information on bacterial community in fog/cloud droplets are scarce, our study provided comprehensive investigation of bacterial community. From the 13 samples collected during 7 cloud episodes, a total of 232148 high quality sequences were obtained after quality filtering and OTUs ranged from 975 to 1258 (Table 2). This value was similar with the previous sequence-based survey of atmospheric bacteria in dust storm (OTUs, 1214) (Katra et al., 2014), and bacteria in rain water in July (OTUs, 1542) (Cho & Jang, 2014). Identification of OTUs at different taxonomic levels yielded 359 species, 411 genera, 152 families, 70 orders, 38 classes and 26 phyla. Across all samples, Proteobacteria was the dominant phylum, followed by Bacteroidetes, Cyanobacteria, Firmicutes, Deinococcus-Thermus, Actinobacteria and Nitrospirae (Figure 1). Bacterial community structure is similar to few other studies explored bacterial diversity in fog/cloud samples, the aforementioned phyla contained a series of genera participate in the atmospheric hydrological cycle (Amato et al., 2007b; Delort et al., 2010). They are predominant taxa in clouds at a high elevation determined by sanger sequencing and tagged pyrosequencing (Bowers et al., 2009), and which are also the typical culturable heterotrophic bacteria from clouds broadly distributed in aquatic and terrestrial habitats (Amato et al., 2005; Kourtev et al., 2011). In the present study, Figure S5 shows the dominant genera collected during cloud process. The predominant genera from Proteobacteria (including *Acinetobacter*, *Stenotrophomonas*, *Pseudomonas*, *Sphingomonas*, *Massilia*, *Delftia*, *Brevundimonas*), Firmicutes (*Bacillus*) and Bacteroidetes (*Empedobacter*) were similar across all samples. The identified genera in cloud water were similar to the limited data descripted microorganisms in fog/cloud water. Fuzzi et al. (1997) investigated bacteria in fog droplets in a highly polluted area and found the predominant genera from *Pseudomonas*, *Bacillus* and *Acinetobacter*. Amato et al (2007b) observed more diverse genera from the phylum of Proteobacteria, Bacteroidetes, Actinobacteria and Firmicutes, which mainly belonging to *Pseudomonas*, *Sphingomonas*, *Staphylococcus*, *Streptomyces* and *Arthrobacter*. Ahern et al (2006) investigated bacterial community

in clouds collected in Scotland and found the dominant species were from *Pseudomonas* and *Acinetobacter*.

Bacterial community function was estimated with PICRUST algorithm. After PICRUST analysis, pathways with participants less than 10% were removed, leaving 225 non-human-gene KEGG pathways. These predominant pathways showed in Figure S6 were mainly related to Amino Acid Metabolism, Carbohydrate Metabolism, Cell Motility, Cellular Processes and Signaling, Energy Metabolism, Enzyme Families, Folding, Sorting and Degradation, Membrane Transport, Nucleotide Metabolism, Nucleotide Metabolism, Replication and Repair, Signal Transduction, Transcription, Translation. Besides the pathways associated with microbial physiological metabolism, we focused on the microbial pathways of metabolic processes in a variety of environments. Fog/cloud droplets contains carbon and nitrogen compounds, which could be available substrate for microbial metabolism in the atmosphere. The predicted function of metabolism was likely attributed to the bacterial gene from the identified taxa (Figure 2). Previous studies have demonstrated that atmospheric bacterial community contained a metabolically diverse group found in a wide range of water/soil habitats. For example, *Acinetobacter*, the most abundant genera widely distributed in land or ocean, was positively associated with the biodegradation, leaching and removal of several organic and inorganic man-made hazardous wastes (Abdelelhaleem, 2003). *Stenotrophomonas* and *Pseudomonas*, positively correlation with carbohydrate metabolism and glycan biosynthesis and metabolism, are well-known for the striking capability to utilize numerous carbon sources, have been widely utilized in the degradation and transformation of complex organic compounds in a wide range of habitats (Boonchan et al., 1998; Stanier et al., 2010). Moreover, predicted functions associated with human disease are especially concerned. For instance, some species from *Acinetobacter*, were positively associated with infection disease (Nemec et al., 2001). *Empedobacter* from Bacteroidetes widely distributed in water habitats, are human clinical origins, certain species from *Empedobacter* are ranked as potential pathogens (Hugo et al., 2005).

In cloud water, a variety of genera adapt to harsh environments were also identified. *Sphingomonas*, the ability to survive in low concentrations of nutrients has been reported, which can metabolize a series of carbon compounds, events toxic compounds (Xu et al., 2006). Similar to *Sphingomonas*, members of *Brevundimonas* are well known for their ability to withstand extreme harsh environment (Kopcakova

et al., 2014). The spore forming bacteria *Bacillus* included in the phylum Firmicutes are commonly airborne bacteria found in bioaerosol, cloud water, rainwater and could survive in cold environment (Despr és et al., 2012). Similar to *Bacillus*, some strains of *Pseudomonads* found in Antarctic environments revealed the cold adaption (Bozal et al., 2003). Certain *Pseudomonads* species found in cloud water were psychrophiles, they grow faster at 5 ℃ than at high temperature (17 ℃ or 27 ℃) (Amato et al., 2007b). Members of *Deinococcus* from Deinococcus-Thermus are available to withstand extreme radiation conditions that could potentially adapt to the cloud environment (Mattimore & Battista, 1996).

Although most the bacterial ecophysiological role in biogeochemical cycles is generally established based on soils and water habitats, information about bacterial activity in cloud water is available. The identification of microorganisms under barren nutrition, low temperatures and radiation environment encountered in clouds is expected since similar bacterial species have been retrieved and proved to be active in harsh environments. Their adaption to the specific environments in fog/cloud water with the potential role in the nucleation, metabolism of organic pollutants, demonstrated the potential importance in participation and influence atmospheric biochemistry cycle.

3.3 Implications in human health and ecosystem

Bacteria in fog/cloud water have been discovered for decades but detailed information on community composition and potential ecophysiological role is severely limited. Bioaerosols in fog/cloud have been complex assemblages of airborne and exogenic microorganisms, likely emissions and resuspension from various terrestrial environments, e.g., soil, water, plants, animals or human beings. In the atmosphere, fog/clouds may be favorable niche for bacteria and these bacteria could thrive and influence cloud processes by acting as cloud condensation nuclei and ice nuclei. Bacteria including pathogenic or beneficial species can also be attached to particles or incorporated into water droplets of fog/clouds. During fog/cloud or rain process, they can be deposited back to land via deposition and possibly induce infections to human health and impose effect on the diversity and function of aquatic and terrestrial ecosystems (Kaushik & Balasubramanian, 2012; Simmons et al., 2001; Va ïilingom et al., 2012) (Figure 3 and Table 3).

Atmospheric bacteria are efficient cloud condensation nuclei and water vapour can be condensed on bacterial cell surface (Mohler et al., 2008). The hygroscopic growth of bacteria below water saturation and supersaturations has been observed for some species, e.g. Bauer et al. (2003) found that *Brevundimonas diminuta* was activated at $<$ 0.1% supersaturation (Bauer et al., 2003). In addition, various strains of *Pseudomonas*, *Rhodococcus* and *Bacillus* found in cloud water samples could produce biosurfactants and act as cloud condensation nuclei (Delort et al., 2010). They may form cloud droplets combined with aerosol particles at lower supersaturations and quickly grow to large size droplets and facilitate rain formation (Mohler et al., 2007). Moreover, *Pseudomonas* could induce ice nucleation at a warmer temperature than usual (Amato et al., 2015). Simulations experiments about cloud forming suggest that *Pseudomonas* was first acted as CCN, then induced freezing and ice nucleation process (Mohler et al., 2008). In addition to *Pseudomonas*, other bacteria from *Acinetobacter*, *Bacillus*, *Flavobacterium*, *Sphingomonas*, and *Stenotrophomonas* sp. (Table 3), were ice nucleation active (Mortazavi et al., 2008). Gaining an understanding of possible role in cloud condensation and ice nucleation processes might open a new sight of bacterial communities influence on meteorology and climate change.

In addition to the bacteria survive in cold environments and act as efficient cloud condensation nuclei or ice nuclei, the presence of microorganisms in fog/cloud may play vital role in atmospheric biochemistry. The detection of bacteria in cloud water associated with biotransformation of organic compounds raised a general understanding of the potential role in atmospheric chemistry. The identified species from the genera of *Rhodococcus*, *Sphingomonas*, *Delftia*, *Comamonas* (Table 3) were mainly participated in the biodegradation of organic compounds. Excessive studies have illustrated their capability of metabolism of hydrocarbon compounds, even toxic pollutants, e.g., aromatic compounds (Bock et al., 1996; Busse et al., 2003; Geng et al., 2009; Goyal & Zylstra, 1996). Two strains from *Stenotrophomonas* (*S. rhizophila*) and *Phyllobacterium* (*P. myrsinacearum*) are typical rhizospheric microorganisms, which were typically dispersed into atmosphere from soil. As plant-associated strains, *S. rhizophila* fulfill plant-protective roles and have been safely applied in biotechnology (Alavi et al., 2013). *P. myrsinacearum* is a predominant rhizospheric bacterium, which has been utilized in plant growth promotion and biological control

of soil-borne diseases due to its capability of azotification (Gonzalezbashan et al., 2000). The methylotrophic bacteria *Methylobacterium* (*M. aquaticum* and *M. adhaesivum*) are typically inhabit in soil and water. Previous studies have demonstrated the carbon fixing function in ecosystem (Gallego et al., 2006; Gallego et al., 2005). Similar to *Methylobacterium*, *Cyanobacterium* sp., widely distributed in soil, water, and various arid environments, have excellent nitrogen and carbon fixing ability (Jha et al., 2004). Cloud water seems to harbor highly diverse bacterial communities in ecosystem, which may be the atmospheric mixing of diverse point sources origins in the rhizosphere, soil and water and possibly participate in the biodegradation of organic compounds in cloud water.

In addition, after sequencing, bacterial genera containing potential pathogens were especially concerned. By blast with the reference pathogen database, sequences high similar with potential pathogens were identified. In the present study, the presence of potential pathogen sequences indicated occasional distribution and dispersion of pathogens in cloud water (Table 3). The identified opportunistic pathogens from *Empedobacter*, e.g., *E. brevis*, can be easily isolated from clinical resources, which may be associated with eye infections (Bottone et al., 1992). Occurrence of *Staphylococcus equorum* in cloud water can be expected since *Staphylococcus* are frequently isolated from airborne samples (Seo et al., 2008). They can reside on the skin and mucous membranes of humans and induce severe infections (Nov ákov á et al., 2006). Similarly, species from *Brevundimonas* (*B. vesicularis* and *B. diminuta*) can induce virulent infections, often associated with nervous system or bacteraemia (Gilad et al., 2009; Han & Andrade, 2005). Besides that, the pathogenic strains from *Acinetobacter* (*A. schindleri*) and *Moraxella* (*M. osloensis*) are associated with skin and wound infections, bacteremia and pneumonia (Banks et al., 2007; Nemec et al., 2001).

Previous studies on potential pathogens are mostly focused on the atmospheric particulate matter ($PM_{2.5}$ and PM_{10}) (Cao et al., 2014; Creamean et al., 2013), rain water (Kaushik & Balasubramanian, 2012; Simmons et al., 2001), and indicated that health risk-related bacteria in atmospheric samples should be concerned. For cloud/fog water, studies of health risks to individuals are typically focused on the chemical characteristic, e.g., the low pH (acid fog) (Hackney et al., 1989), PAH (Ehrenhauser et al., 2012), etc. Limited literature discussed the microorganism in fog/cloud water suggested potential pathogens in fog/cloud water (Va ïilingom et al.,

2012), they find potential plant pathogens such as *Pseudomonas syringae* and *Xanthomonas campestris* and suggest these living plant pathogens could then infect new hosts through precipitation. Possibly, greater survival of human pathogens may be supported in the atmosphere. Fog/cloud and rain process are part of atmospheric life cycle and dispersal pathway for some pathogenic bacteria. Studies of the airborne dispersal of pathogenic bacteria, e.g., *Neisseria meningitides*, *Staphylococcus aureus* from dust samples from Kuwait and *Pseudomonas aeruginosa* in USA Virgin Islands have indicated the spread of specific human and plant diseases over long term transport in atmosphere (Griffin, 2007; Griffin et al., 2003; Griffin et al., 2006). However, detailed health risk-oriented studies induced by pathogenic microorganisms should be deeply conducted and prudently assessed. Further study depending on the culture-dependent method and biochemical experiments will perform to check the pathogenicity.

3.4 Disparity between polluted and non-polluted cloud episodes

To distinguish indicator species within the polluted and non-polluted cloud episodes, LEfSe is performed, which showed statistically significant differences. A total of 70 bacterial groups were distinct using the default logarithmic (LDA) value of 2. Cladograms show taxa with LDA values higher than 3.5 for clarity (Figure 4). Consequently, 8 and 19 represent bacterial taxa in polluted and non-polluted cloud episodes were detected.

In polluted cloud episodes, most enriched bacteria were ranked as opportunistic pathogens, such as Proteobacteria, Gammaproteobacteria, Xanthomonadales, Xanthomonadaceae, Stenotrophomonas, Moraxellaceae and *Acinetobacter*. Like Proteobacteria, the Gram-negative Gammaproteobacteria contains a series of ecologically and medically important bacteria, e.g., the pathogenic Enterobacteriaceae, Vibrionaceae and Pseudomonadaceae. The Xanthomonadales from this phylum has been reported to cause disease in plants (Saddler & Bradbury, 2005). Certain species from *Stenotrophomonas* (Gammaproteobacteria, Xanthomonadales) are associated with multiple human infections. *Moraxella* form Moraxellaceae (Gammaproteobacteria) has been reported to associate with septic arthritis of the ankle (Banks et al., 2007). As previous mentioned, species from the genus of *Acinetobacter* are opportunistic pathogens and cause severe clinical infections

(Nemec et al., 2001).

In comparison, the majority of indicator species in the non-polluted samples are from Bacteroidetes, Firmicutes, Betaproteobacteria and Deinococcus-Thermus. An important biomarker from Bacteroidetes is Flavobacteriia, relative study has illustrated the marine sources for Flavobacteria. Most of Flavobacteria sequences searched by NCBI blast are mainly from marine sources, i.e., algae, oysters and sea cucumbers (Cho & Jang, 2014). The genera *Empedobacter* was abundant across all samples, which are included in the family Flavobacteriaceae. As mentioned above, *Empedobacter* **(**Firmicutes) are potential pathogens and resistant to a wide range of antimicrobials (Hugo et al., 2005). Clostridiales (Clostridia) and *Bacillus* (Bacillaceae) are two represent biomarkers from Firmicutes identified in the non-polluted cloud water samples. As ubiquitous in nature, these two groups contain some medically significant species (Miller et al., 2001; Makino & Cheun, 2003). Moreover, their specific physiological characteristics (produce a variety of enzymes and metabolites) and excellent ability to decomposition of organic matter have made the widely utilization in biotechnology and fermentation industry (Doi et al., 1992; Łoś, et al., 2010). Members of Oxalobacteraceae (Betaproteobacteria) are rhizosphere microorganisms involved in the biological nitrogen fixation (Donn et al., 2015). The Burkholderiales (Betaproteobacteria) commonly found in water and soil are involved in the biodegradation of various aromatic compounds (Pérez-Pantoja et al., 2012). Deinococci, from the phylum of Deinococcus-Thermus could resistant to extreme radiation and survive in extremes of heat and cold (Griffiths & Gupta, 2007).

By comparison, potential pathogens were significant groups in the polluted samples, whereas a diverse ecological function group was identified in the non-polluted samples originated from a wide range of habitat. Ecologically meaningful distinguish of bacterial groups under polluted and non-polluted conditions is essential for understanding the structure and function, and which provide a general understanding of bacterial metabolism in cloud water.

3.5 Environmental factors shaping bacterial community structure

To clarity the vital environmental factor in shaping bacterial community structure, RDA was performed to discern the genus-level structure with the selected environmental factors (Figure 5). The first two *axes* explained 65.9% of the

accumulated variance in the species-environment relation. *Interset correlations* showed major ions and $PM_{2.5}$ was the most important environmental variables structuring the bacterial community (axis 1, major, -0.436; $PM_{2.5}$, -0.367); in turn, wind speed and temperature registered the high value for axis 2 (wind speed, -0.509; temperature, -0.494) (Table S1).

Cumulative fit indicated that the predominant genera affiliated with *Acinetobacter*, *Bacillus*, *Corynebacterium*, *Phyllobacterium*, *Pseudoalteromonas* and *Rhodococcus* displayed strong correlations with axis 1. *Empedobacter*, *Hydrotalea*, *Paracoccus*, *Pelomonas*, *Pseudomonas* and *Stenotrophomonas* were notable genera highly correlations with axis 2 (Table S2). As aforementioned in section 3.3, the above bacteria were metabolically diverse groups found in various habitats and certain genera included potential pathogens.

Of the environmental characteristics measured, major ions in cloud water and atmospheric $PM_{2.5}$ were the best predictors of diversity variability of bacterial community structure. These two parameters were strongly correlated with represent bacterial genera. As indicated in Figure S1, significant positive correlation was observed between major ions (x) and $PM_{2.5}$ (y), $y=0.00477x+5.324$, $p<0.01$, $R^2=0.757$. Relevant studies suggest that bacterial community was highly variable under $PM_{2.5}$ mass concentration (Cao et al., 2014). Statistical analysis, e.g., correlation or multiple linear regression, indicated that $PM_{2.5}$ exhibited a negative correlation with airborne bacteria in haze days (Gandolfi et al., 2015; Gao et al., 2015), whereas in another study, spearman correlation analysis showed $PM_{2.5}$ exhibited a significant positive correlations with the airborne microbe concentration (Dong et al., 2016). Possibly, the inorganic and organic compounds in particulate matter ($PM_{2.5}$) can be available nutrients for microbial growth in air. However, aggregation of harmful substances such as heavy metals, polycyclic aromatic hydrocarbons would be toxic for bacteria under high $PM_{2.5}$ mass concentration.

During cloud process, most atmospheric particles (including $PM_{2.5}$) are scavenged in cloud water. In polluted air, high $PM_{2.5}$ concentration resulted in the elevated water soluble inorganic ions in cloud water droplets. Therefore, similar trends were observed between major ions and $PM_{2.5}$ concentration. In cloud water, the water soluble major ions and the microorganisms co-exist in the same microenvironment. Major ions in cloud droplets could provide available nutrition for bacterial growth and

duplication. Previous study has suggested that these nutrients are related to the distribution of bacteria in water habitats (Newton et al., 2011). Meanwhile, PICRUST analysis in section 3.2 discovered a series of metabolic pathway involved in the bacterial basic physiological activities and carbon, nitrogen, sulfur metabolism (Figure S6). For example, the sulfate, nitrate and ammonium can be available substrates for bacterial growth, magnesium and calcium are involved in a series of physiological activities (e.g., signal regulation, transmembrane transport) (Fagerbakke, et al., 1999; Fiermonte, et al., 2004; Michiels et al., 2002). Therefore, major ions were important environmental factor shaping community structure in cloud water. $PM_{2.5}$ played an indirect role by influencing the concentration of major ions in water droplets.

The identified taxa either from polluted or non-polluted samples were typically found in soil, water, plant or human beings. These bacterial groups aerosolized and dispersed into atmosphere either from local regional emissions or long-term transport. Source tracking analysis by backward trajectory indicated that air mass of polluted cloud episodes came largely from northern and western China, moved east through Shanxi, Henan, Hebei province to the study area, or from outer Mongolia, crossing the Jingjinji area to Mt.Tai (Figure 6). The passed areas were notable as heavy industry region with frequent coal mining activities and serious pollution. Moreover, the large population and agricultural activities resulted in the numerous pathogenic microorganisms from human or animal fecal dispersed in the atmosphere. In contrary, air mass of non-polluted cloud episodes originated mostly from the southern China, and the passed region were rich of water resources, e.g., Dongting Lake, Huaihe river, Yangtze river etc. The marine sources bacteria (Flavobacteria, significant biomarker in non-polluted cloud water samples by LefSe, Figure 4) dispersed in the atmosphere typically derived from the evaporation of lakes and rivers water. These bacteria mainly originated from sea-air interactions and airborne marine bacteria can be transported to inland through long-term transport.

In the sampling site (the sumit of Mt.Tai, 1534 m a.s.l), local anthropogenic pollution might be minimized and air pollution is mainly influenced by long term transport. Wind rose diagram suggest the predominant west wind during polluted cloud episodes and wind speeds ranged 1.2-1.4 m/s, whereas in non-polluted cloud episodes it was mainly southwest wind with higher wind speed (2.4-3 m/s) (Figure 7). Wind direction and speed are important meteorological factors influencing fog/cloud formation (Fu et

al., 2014). Recent studies also indicate the variation of bacterial concentration and community structure conducted by wind (Evans et al., 2006; Jones & Harrison, 2004). In addition, wind and $PM_{2.5}$ distribution graph (Figure 7C) indicates that during the whole sampling time (from 24 July to 23 August 2014) $PM_{2.5}$ concentration was high under lower wind speed, whereas lower $PM_{2.5}$ was observed when wind speed was high. Air mass from the contaminated area through long-term transport combined with lower wind speeds, largely reduce the diffusion rate of pollutants and thus lead to the sustained high $PM_{2.5}$ during polluted cloud episodes. In polluted air, the soluble composition in $PM_{2.5}$ was accumulated in atmosphere and could be dissolved in cloud water droplets during wet deposition. Therefore, the concentration of water soluble ions increased under high $PM_{2.5}$ concentration, which has directly influence on microbial community. Whereas in the non-polluted cloud episodes, higher wind speed was beneficial to the diffusion of pollutants and resulted in the lower $PM_{2.5}$ mass concentration, which finally led to the significant discrepancy of bacterial community structure. However, further research still needed to address the detailed interaction between bacterial community and environmental factors, and to understand the mechanism that how chemical composition influence microbial community.

4 Conclusion

The composition and potential function of microbial communities in atmospheric water phase (fog and clouds) remained rarely studied. Using the 16S rRNA gene sequencing, this work provided a thorough investigation on bacterial ecological diversity under polluted and non-polluted cloud episodes and revealed a highly diverse bacterial community harbored in cloud water. Correlation analysis for the predominant genera and PICRUST function predication enhanced the understanding of the distribution of bacteria and their potential involvements in atmosphere, ecosystem and human health. The identification of bacteria survive in barren nutrition, low temperatures and radiation environment encountered in fog/cloud water demonstrated bacterial active in harsh atmospheric environments. They may act as efficient cloud condensation nuclei or ice nuclei, associate with biogeochemical cycling (nitrogen/carbon cycling) and microbial degradation of organic compounds in fog/clouds and spreading of specific human, animal and plant diseases by potential pathogens. Moreover, community disparity between polluted and non-polluted cloud

episodes suggested that major ions in cloud water and atmospheric $PM_{2.5}$ seem to be pivotal variables in shaping bacterial community. $PM_{2.5}$ had potential impact on bacterial community structure by influencing major ions in water droplets, which is likely to provide a more comprehensive understanding of the atmospheric microbial biodiversity under environmental stress. These results provide a basic understanding of mechanism of bacterial community response and metabolism in polluted weather for further study.

Acknowledgement

This work was supported by National Natural Science Foundation of China (41605113, 41375126), Taishan Scholar Grand (ts20120552), China Postdoctoral Science Foundation (No. 2015M582095).

Reference

Abdelelhaleem, D.: *Acinetobacter*: environmental and biotechnological applications, Afr. J. Biotechnol., 4, 71-74, 2003.

Adhikari, A., Reponen, T., Grinshpun, S.A., Dainius. Martuzevicius, LeMasters, G.: Correlation of ambient inhalable bioaerosols with particulate matter and ozone: A two-year study, Environ. Pollut., 140, 16-28, 2006.

Ahern, H.E., Walsh, K.A., Hill, T.C.J., Moffett, B.F.: Fluorescent pseudomonads isolated from Hebridean cloud and rain water produce biosurfactants but do not cause ice nucleation, Biogeosciences. Dis., 3, 985-9, 2006.

Aikawa, M., Hiraki, T., Shoga, M., Tamaki, M.: Fog and precipitation chemistry at Mt. Rokko in Kobe, April 1997-March 1998, Water. Air.Soil. Poll., 130, 1517-1517, 2001.

Alavi, P., Starcher, M.R., Thallinger, G.G., Zachow, C., Müller, H., Berg, G.: *Stenotrophomonas* comparative genomics reveals genes and functions that differentiate beneficial and pathogenic bacteria, BMC. Genomics., 15, 1-15, 2013.

Amato, P., Demeer, F., Melaouhi, A., Fontanella, S., Martin-Biesse, A.-S., Sancelme, M., et al.: A fate for organic acids, formaldehyde and methanol in cloud water:their biotransformation by micro-organisms, Atmos. Chem. Phys., 7, 4159-4169, 2007a.

Amato, P., Joly, M., Schaupp, C., Attard, E., Möhler, O., Morris, C.E., et al.: Survival and ice nucleation activity of bacteria as aerosols in a cloud simulation chamber, Atmos. Chem. Phys., 15, 6455-6465, 2015.

Amato, P., Ménager, M., Sancelme, M., Laj, P., Mailhot, G., Delort, A.-M.: Microbial population in cloud water at the Puy de Dôme: Implications for the chemistry of clouds, Atmos. Environ., 39, 4143-4153, 2005.

Amato, P., Parazols, M., Sancelme, M., Laj, P., Mailhot, G., Delort, A.M.: Microorganisms isolated from the water phase of tropospheric clouds at the Puy de Dome: major groups and growth abilities at low temperatures, FEMS. Microbiol. Ecol., 59, 242-254, 2007b.

Amato, P., Parazols, M., Sancelme, M., Mailhot, G., Laj, P., Delort, A.-M.: An important oceanic source of micro-organisms for cloud water at the Puy de Dôme (France), Atmos. Environ., 41, 8253-8263, 2007c.

Anderson, M.J., Willis, T.J.: Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology, Ecology., 84, 511-525, 2008.

Badarinath, K.V.S., Latha, K.M., Chand, T.R.K., Reddy, R.R., Gopal, K.R., Reddy, L.S.S., et al.: Black carbon aerosols and gaseous pollutants in an urban area in North India during a fog period, Atmos. Res., 85, 209-216, 2007.

Banks, L.N., Kurdy, N.M., Hassan, I., Aster, A.S.: Septic arthritis of the ankle due to *Moraxella osloensis*, Foot. Ankle. Surg., 13, 189-191, 2007.

Bauer, H., Giebl, H., Hitzenberger, R., Anne, K.G., Georg, R., Franziska, Z., et al.: Airborne bacteria as cloud condensation nuclei, J. Geophys. Res., 108, 1919-1964, 2003.

Bell, S.G., Wong, L.L.: P450 enzymes from the bacterium *Novosphingobium aromaticivorans*, Biochem. Bioph. Res. Co., 360, 666-672, 2007.

Ben Abdeljalil, N.O., Vallance, J.: Biocontrol of rhizoctonia root rot in tomato and enhancement of plant growth using Rhizobacteria naturally associated to tomato, J. Plant. Pathol. Microbiol., 7, 10.4172/2157-7471.1000356, 2016.

Bertolini, V., Gandolfi, I., Ambrosini, R., Bestetti, G., Innocente, E., Rampazzo, G., et al.: Temporal variability and effect of environmental variables on airborne bacterial communities in an urban area of Northern Italy, Appl. Microbiol. Biotechnol., 97, 6561-6570, 10.1007/s00253-012-4450-0, 2013.

Bibby, K., Viau, E., Peccia, J.: Bibby K, Viau E, Peccia J.. Pyrosequencing of the 16S rRNA gene to reveal bacterial pathogen diversity in biosolids. Water Res 44: 4252-4260, Water. Res., 44, 4252-4260, 2010.

Bock, C., Kroppenstedt, R.M., Diekmann, H.: Degradation and bioconversion of aliphatic and aromatic hydrocarbons by *Rhodococcus ruber* 219, Appl. Microbiol. Biot., 45, 408-410, 1996.

Boonchan, S., Britz, M.L., Stanley, G.A.: Surfactant-enhanced biodegradation of high molecular weight polycyclic aromatic hydrocarbons by *Stenotrophomonas maltophilia*, Biotechnol. Bioeng., 59, 482-494, 1998.

Boris, A.J., Lee, T., Park, T., Choi, J., Seo, S., Collett, J.L.: Fog composition at Baengnyeong Island in the Eastern Yellow Sea: detecting markers of aqueous atmospheric oxidations, Atmos. Chem. Phys., 15, 24871-24908, 2015.

Bottone, E.J., Madayag, R.M., Qureshi, M.N.: Acanthamoeba keratitis: synergy between amebic and bacterial cocontaminants in contact lens care systems as a prelude to infection, J. Clin. Microbiol., 30, 2447-2450, 1992.

Bowers, R.M., Clements, N., Emerson, J.B., Wiedinmyer, C., Hannigan, M.P., Fierer, N.: Seasonal variability in bacterial and fungal diversity of the near-surface atmosphere, Environ. Sci. Technol., 47, 12097-12106, 2013.

Bowers, R.M., Lauber, C.L., Wiedinmyer, C., Hamady, M., Hallar, A.G., Fall, R., et al.: Characterization of airborne microbial communities at a high-elevation site and their potential to act as atmospheric ice nuclei, Appl. Environ. Microbiol., 75, 5121-5130, 2009.

Bozal, N., Montes, M.J., Tudela, E., Guinea, J.: Characterization of several Psychrobacter strains isolated from Antarctic environments and description of *Psychrobacter luti* sp. nov. and *Psychrobacter fozii* sp. nov, Int. J. Syst. Evol. Micr., 53, 1093-1100, 2003.

Burrows, S.M., Elbert, W., Lawrence, M.G., Pöschl, U.: Bacteria in the global atmosphere-part 1: review and synthesis of literature data for different ecosystems, Atmos. Chem. Phys., 9, 9263-9280, 2009.

Busse, H., Denner, E., Buczolits, S., Salkinoja-Salonen, M., Bennasar, A., Kämpfer, P.: *Sphingomonas aurantiaca* sp. nov., *Sphingomonas aerolata* sp. nov. and *Sphingomonas faeni* sp. nov., air- and dustborne and Antarctic, orange-pigmented, psychrotolerant bacteria, and emended description of the genus *Sphingomonas*, Int. J. Syst. Evol. Micr.,

53, 1253-1260, 2003.

Cao, C., Jiang, W., Wang, B., Fang, J., Lang, J., Tian, G., et al.: Inhalable microorganisms in Beijing's $PM_{2.5}$ and PM_{10} pollutants during a severe smog event, Environ. Sci. Technol., 48, 1499-1507, 2014.

Cho, B.C., Jang, G.I.: Active and diverse rainwater bacteria collected at an inland site in spring and summer 2011, Atmos. Environ., 94, 409-416, 2014.

Clarke, K.R.: Non-parametric multivariate analyses of changes in community structure, Austral. Ecol., 18, 117-143, 1993.

Corrigan, A., de Leeuw, M., Penaud-Frezet, S., Dimova, D., Murphy, R.A.: Phylogenetic and functional alterations in bacterial community compositions in broiler ceca as a result of mannan oligosaccharide supplementation, Appl. Environ. Microbiol., 81, 3460-3470, 2015.

Creamean, J.M., Suski, K.J., Rosenfeld, D., Cazorla, A., Demott, P.J., Sullivan, R.C., et al.: Dust and biological aerosols from the Sahara and Asia influence precipitation in the western U.S, Science, 339, 1572-8, 2013.

Delort, A.M., Vaïtilingom, M., Amato, P., Sancelme, M., Parazols, M., Mailhot, G., et al.: A short overview of the microbial population in clouds: Potential roles in atmospheric chemistry and nucleation processes, Atmos. Res., 98, 249-260, 2010.

Després, V.R., Alex Huffman, J., Burrows, S.M., Hoose, C., Safatov, A.S., Buryak, G., et al.: Primary biological aerosol particles in the atmosphere: a review, Tellus. B., 64, 145-153, 2012.

Doi R H, Mcgloughlin M. Biology of Bacilli. Applications to industry. Butterworth-Heinemann, 1992:1-359.

Dong, L., Qi, J., Shao, C., Zhong, X., Gao, D., Cao, W., et al.: Concentration and size distribution of total airborne microbes in hazy and foggy weather, Sci. Total. Environ., 541, 1011-1018, 2016.

Donn, S., Kirkegaard, J.A., Perera, G., Richardson, A.E., Watt, M.: Evolution of bacterial communities in the wheat crop rhizosphere, Environ. Microbiol., 17, 610–621, 2015.

Ehrenhauser, F.S., Khadapkar, K., Wang, Y., Hutchings, J.W., Delhomme, O., Kommalapati, R.R., et al.: Processing of atmospheric polycyclic aromatic hydrocarbons by fog in an urban environment, J. Environ. Monitor., 14, 2566-2579, 2012.

Evans, C.A., Coombes, P.J., Dunstan, R.H.: Wind, rain and bacteria: The effect of weather on the microbial composition of roof-harvested rainwater, Water. Res., 40, 37-44, 2006.

Fagerbakke K M, Norland S, Heldal M. The inorganic ion content of native aquatic bacteria. Can. J. Microbiol., 45, 304, 1999.

Fernández-González, R., Yebra-Pimentel, I., Martínez-Carballo, E., Simal-Gándara, J., Pontevedra-Pombal, X.: Atmospheric pollutants in fog and rain events at the northwestern mountains of the Iberian Peninsula, Sci. Total. Environ., 497–498, 188-199, 2014.

Fiermonte G, De L F, Todisco S, et al. Identification of the mitochondrial ATP-Mg/Pi transporter. Bacterial expression, reconstitution, functional characterization, and tissue distribution. J. Biol. Chem., 279, 30722-30730, 2004.

Fu, G.Q., Xu, W.Y., Yang, R.F., Li, J.B., Zhao, C.S.: The distribution and trends of fog and haze in the North China Plain over the past 30 years, Atmos. Chem. Phys., 14, 11949-11958, 2014.

Fuzzi, S., Mandrioli, P., Perfetto, A.: Fog droplets-an atmospheric source of secondary biological aerosol particles, Atmos. Environ., 31, 287-290, 1997.

Gallego, V., García, M.T., Ventosa, A.: *Methylobacterium adhaesivum* sp. nov., a methylotrophic bacterium isolated from drinking water, Int. J. Syst. Evol. Micr., 56, 339-342, 2006.

Gallego, V., García, M.T., Ventosa, A.: Methylobacterium hispanicum sp. nov. and Methylobacterium aquaticum sp. nov., isolated from drinking water, Int. J. Syst. Evol.

Micr., 55, 281-287, 2005.

Gandhi, P.M., Narayanan, K.B., Naik, P.R., Sakthivel, N.: Characterization of *chryseobacterium aquaticum* strain PUPC1 producing a novel antifungal protease from rice rhizosphere soil, J. Microbiol. Biotechn., 19, 99-107, 2009.

Gandolfi, I., Bertolini, V., Bestetti, G., Ambrosini, R., Innocente, E., Rampazzo, G., et al.: Spatio-temporal variability of airborne bacterial communities and their correlation with particulate matter chemical composition across two urban areas, Appl. Microbiol. Biot., 99, 4867-4877, 2015.

Gao, M., Qiu, T., Jia, R., Han, M., Song, Y., Wang, X.: Concentration and size distribution of viable bioaerosols during non-haze and haze days in Beijing, Environ. Sci. Pollut. R., 22, 4359-4368, 2015.

Gao, M., Yan, X., Qiu, T., Han, M., Wang, X.: Variation of correlations between factors and culturable airborne bacteria and fungi, Atmos. Environ., 128, 10-19, 2016.

Geng, L., Ming, C., Liang, Q., Wei, L., Wei, Z., Ping, S., et al.: Functional analysis of a putative regulatory gene, tadR , involved in aniline degradation in *Delftia tsuruhatensis* AD9, Arch. Microbiol., 191, 603-614, 2009.

Gilad, J., Borer, A., Peled, N., Riesenberg, K., Tager, S., Appelbaum, A., et al.: Hospital-acquired brevundimonas vesicularis septicaemia following open-heart surgery: case report and literature review, Infect. Dis., 32, 90-91, 2009.

Glaeser, S.P., Kämpfer, P. The Family Sphingomonadaceae. Springer Berlin Heidelberg, 2014.

Gonzalezbashan, L.E., Lebsky, V.K., Hernandez, J.P., Bustillos, J.J., Bashan, Y.: Changes in the metabolism of the microalga *Chlorella vulgaris* when coimmobilized in alginate with the nitrogen-fixing *Phyllobacterium myrsinacearum*, Can. J. Microbiol., 46, 653-659, 2000.

Gopalakrishnan, S., Srinivas, V., Prakash, B., Sathya, A., Vijayabharathi, R.: Plant growth-promoting traits of Pseudomonas geniculata isolated from chickpea nodules, 3 Biotech, 5, 653-661, 2015.

Goyal, A.K., Zylstra, G.J.: Molecular cloning of novel genes for polycyclic aromatic hydrocarbon degradation from Comamonas testosteroni GZ39, Appl. Environ. Microbiol., 62, 230-236, 1996.

Griffin, D.W.: Atmospheric movement of microorganisms in clouds of desert dust and implications for human health, Clin. Microbiol. Rev., 20, 459-477, 2007.

Griffin, D.W., Kellogg, C.A., Garrison, V.H., Lisle, J.T., Borden, T.C., Shinn, E.A.: Atmospheric microbiology in the northern Caribbean during African dust events, Aerobiologia, 19, 143-157, 2003.

Griffin, D.W., Westphal, D.L., Gray, M.A.: Airborne microorganisms in the African desert dust corridor over the mid-Atlantic ridge, Ocean Drilling Program, Leg 209, Aerobiologia, 22, 211-226, 2006.

Griffiths, E., Gupta, R.S.: Identification of signature proteins that are distinctive of the Deinococcus-Thermus phylum, Int. Microbiol., 10, 201-208, 2007.

Guo, J., Wang, Y., Shen, X., Wang, Z., Lee, T., Wang, X., et al.: Characterization of cloud water chemistry at Mount Tai, China: Seasonal variation, anthropogenic impact, and cloud processing, Atmos. Environ., 60, 467-476, 2012a.

Guo, F., Zhang, T.: Profiling bulking and foaming bacteria in activated sludge by high throughput sequencing, Water. Res., 46, 2772-2782, 2012b.

Hackney, J.D., Linn, W.S., Avol, E.L.: Acid fog: effects on respiratory function and symptoms in healthy and asthmatic volunteers, Environ. Health Perspect., 79, 159-162, 1989.

Han, X.Y., Andrade, R.A.: *Brevundimonas diminuta* infections and its resistance to fluoroquinolones, J. Antimicrob. Chemoth., 55, 853-859, 2005.

Hauser, E., Kämpfer, P., Busse, H.J.: *Pseudomonas psychrotolerans* sp. nov, Int. J. Syst. Evol. Micr., 54, 1633-1637, 2004.

Hill, K.A., Shepson, P.B., Galbavy, E.S., Anastasio, C., Kourtev, P.S., Konopka, A., et al.: Processing of atmospheric nitrogen by clouds above a forest environment, J. Geophys. Res., 112, 2864-2864, 2007.

Huang, K., Zhuang, G., Wang, Q., Fu, J.S., Lin, Y., Liu, T., et al.: Extreme haze pollution in Beijing during January 2013: chemical characteristics, formation mechanism and role of fog processing, Atmos. Chem. Phys. Discuss., 2014, 7517-7556, 2014.

Hugo, C.J., Bruun, B., Jooste, P.J.: The Genera Empedobacter and Myroides, in, ed., edited by, Springer New York, 630-637, 2005.

Jha, M.N., Prasad, A.N., Misra, S.K.: Influence of source of organics and soil organic matter content on cyanobacterial nitrogen fixation and distributional pattern under different water regimes, World. J. Microb. Biot., 20, 673-677, 2004.

Joly, M., Attard, E., Sancelme, M., Deguillaume, L., Guilbaud, C., Morris, C.E., et al.: Ice nucleation activity of bacteria isolated from cloud water, Atmos. Environ., 70, 392-400, 2013.

Jones, A.M., Harrison, R.M.: The effects of meteorological factors on atmospheric bioaerosol concentrations-a review, Sci. Total. Environ., 326, 151-180, 2004.

Kämpfer, P., Lodders, N., Huber, B., Falsen, E., Busse, H.J.: *Deinococcus aquatilis* sp. nov., isolated from water, Int. J. Syst. Evol. Micr., 58, 2803-2806, 2009.

Katra, I., Arotsker, L., Krasnov, H., Zaritsky, A., Kushmaro, A., Ben-Dov, E.: Richness and diversity in dust stormborne biomes at the southeast mediterranean, Scientific reports, 4, 5265, 10.1038/srep05265, 2014.

Kaushik, R., Balasubramanian, R.: Assessment of bacterial pathogens in fresh rainwater and airborne particulate matter using Real-Time PCR, Atmos. Environ., 46, 131-139, 2012.

Kindt, R., Coe, R.: Tree diversity analysis: a manual and software for common statistical methods for ecological and biodiversity studies, J. Am. Vet. Med. A., 235, 368-374, 2005.

Kopcakova, A., Stramova, Z., Kvasnova, S., Godany, A., Perhacova, Z., Pristas, P.: Need for database extension for reliable identification of bacteria from extreme environments using MALDI TOF mass spectrometry, Chem. Pap., 68, 1435-1442, 2014.

Kourtev, P.S., Hill, K.A., Shepson, P.B., Konopka, A.: Atmospheric cloud water contains a diverse bacterial community, Atmos. Environ., 45, 5399-5405, 2011.

Kuczynski, J., Stombaugh, J., Walters, W.A., González, A., Caporaso, J.G., Knight, R.: Using QIIME to analyze 16S rRNA gene sequences from microbial communities:Current Protocols in Microbiology, Curr. Protoc. Bioinformatics., 10, 1-28, 2011.

Langille, M.G., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., et al.: Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences, Nat. Biotechnol., 31, 814-821, 2013.

Liu, X.H., Wai, K.M., Wang, Y., Zhou, J., Li, P.H., Guo, J., et al.: Evaluation of trace elements contamination in cloud/fog water at an elevated mountain site in Northern China, Chemosphere, 88, 531-41, 2012.

Liu, Y., Wang, L., Huang, K., Wang, W., Nie, X., Jiang, Y., et al.: Physiological and Biochemical Characterization of a Novel Nicotine-Degrading Bacterium Pseudomonas geniculata N1, Plos. One., 9, 393-418, 2014.

Łoś, M, Kuzio J, Mcconnell M R, et al. Lysogenic conversion in bacteria of importance to the food industry. 2010.

Luo, G., Angelidaki, I.: Analysis of bacterial communities and bacterial pathogens in a biogas plant by the combination of ethidium monoazide, PCR and Ion Torrent sequencing, Water. Res., 60, 156-163, 2014.

Madsen, A.M., Zervas, A., Tendal, K., Nielsen, J.L.: Microbial diversity in bioaerosol samples causing ODTS compared to reference bioaerosol samples as measured using Illumina sequencing and MALDI-TOF, Environ. Res., 140, 255-267, 2015.

Makino, S.-i., Cheun, H.-i.: Application of the real-time PCR for the detection of airborne microbial pathogens in reference to the anthrax spores, J. Microbiol. Meth., 53, 141-147, 2003.

Masoud, W., Takamiya, M., Vogensen, F.K., Lillevang, S., Al-Soud, W.A., Sørensen, S.J., et al.: Characterization of bacterial populations in Danish raw milk cheeses made with different starter cultures by denaturating gradient gel electrophoresis and pyrosequencing, Int. Dairy. J., 21, 142-148, 2011.

Mattimore, V., Battista, J.R.: Radioresistance of *Deinococcus radiodurans*: functions necessary to survive ionizing radiation are also necessary to survive prolonged desiccation, J. Bacteriol., 178, 633-637, 1996.

McIlroy, S.J., Nielsen, P.H.: The Family Saprospiraceae, in: The Prokaryotes-Other Major Lineages of Bacteria and the Archaea, ed., edited by, E. Rosenberg et al. (eds.), 863-889, 2014.

Meziane, H., Sluis, I.V.D., Loon, L.C.V., Höfte, M., Bakker, P.A.H.M.: Determinants of *Pseudomonas putida* WCS358 involved in inducing systemic resistance in plants, Mol. Plant. Pathol., 6, 177–185, 2005.

Michiels, J., Xi, C., Verhaert, J., Vanderleyden, J.: The functions of Ca^{2+} in bacteria: a role for EF-hand proteins?, Trends. Microbiol., 10, 87-93, 2002.

Miller, D.L., Brazer, S., Murdoch, D., Reller, L.B., Corey, G.R.: Significance of *Clostridium tertium* bacteremia in neutropenic and nonneutropenic patients: review of 32 cases, Clin. Infect. Dis., 32, 975-978, 2001.

Mohan, M., Payra, S.: Influence of aerosol spectrum and air pollutants on fog formation in urban environment of megacity Delhi, India, Environ. Monit. Assess., 151, 265-277, 2009.

Mohler, O., Demott, P.J., Vali, G., Levin, Z.: Microbiology and atmospheric processes: the role of biological particles in cloud physics, Biogeosciences, 4, 1059-1071, 2007.

Mohler, O., Georgakopoulos, D.G., Morris, C.E., Benz, S., Ebert, V., Hunsmann, S., et al.: Aerosol–cloud–precipitation interactions. Part 1. The nature and sources of cloud-active aerosols, Earth-Sci. Rev., 89, 13-41, 2008.

Mortazavi, R., Hayes, C.T., Ariya, P.A.: Ice nucleation activity of bacteria isolated from snow compared with organic and inorganic substrates, Environ. Chem., 5, 373-381, 2008.

Mortazavi, R., Attiya, S., Ariya, P.A.: Arctic microbial and next-generation sequencing approach for bacteria in snow and frost flowers: selected identification, abundance and freezing nucleation, Atmos. Chem. Phys., 15, 6183-6204, 2015.

Nakamura, S., Sakamoto, Y., Ishiyama, M., Tanaka, D., Kunii, K., Kubo, K., et al.: Characterization of two oil-degrading bacterial groups in the Nakhodka oil spill, Int. Biodeter. Biodegr., 60, 202-207, 2007.

Nemec, A., De Baere, T., Tjernberg, I., Vaneechoutte, M., van der Reijden, T.J., Dijkshoorn, L.: *Acinetobacter ursingii* sp. nov. and *Acinetobacter schindleri* sp. nov., isolated from human clinical specimens, Int. J. Syst. Evol. Micr., 51, 1891-1899, 2001.

Newton, R.J., Jones, S.E., Alexander, E., Mcmahon, K.D., Stefan, B.: A guide to the natural history of freshwater lake bacteria, Microbiol. Molecul. Biol. R., 75, 14-49, 2011.

Nováková, D., Sedlácek, I., Pantůcek, R., Stetina, V., Svec, P., Petrás, P.: *Staphylococcus equorum* and *Staphylococcus succinus* isolated from human clinical specimens, J. Med. Microbiol., 55, 523-528, 2006.

Pérez-Pantoja, D., Donoso, R., Agulló, L., Córdova, M., Seeger, M., Pieper, D.H., et al.: Genomic analysis of the potential for aromatic compounds biodegradation in Burkholderiales, Environ. Microbiol., 14, 1091–1117, 2012.

Poeschl, U.: Atmospheric aerosols: composition, transformation, climate and health effects, Angew. Chem. Int. Edit., 37, 7520-7540, 2006.

Ponder, M.A., Gilmour, S.J., Bergholz, P.W., Mindock, C.A., Hollingsworth, R., Thomashow, M.F.,

et al.: Characterization of potential stress responses in ancient Siberian permafrost psychroactive bacteria, FEMS. Microbiol. Ecol., 53, 103-115, 2005.

Reardon, K.F., Mosteller, D.C., Bull Rogers, J.D.: Biodegradation kinetics of benzene, toluene, and phenol as single and mixed substrates for *Pseudomonas putida* F1, Biotechnol. Bioeng., 69, 385-400, 2000.

Saddler, G.S., Bradbury, J.F.: Xanthomonadales ord. nov, in: Bergey's Manual® of Systematic Bacteriology, ed., edited by, Springer US, 63-122, 2005.

Schloss, P., Westcott, S., Ryabin, T., Hall, J., Hartmann, M., Hollister, E., et al.: Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities, Appl. Environ. Microbiol., 75, 7537-7541, 2009.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., et al.: Metagenomic biomarker discovery and explanation, Genome. Biol., 12, 1-18, 2011.

Seo, Y.S., Lee, D.Y., Rayamahji, N., Mi, L.K., Han, S.Y.: Biofilm-forming associated genotypic and phenotypic characteristics of Staphylococcus spp. isolated from animals and air, Res. Vet. Sci., 85, 433-438, 2008.

Simmons, G., Hope, V., Lewis, G., Whitmore, J., Gao, W.: Contamination of potable roof-collected rainwater in Auckland, New Zealand, Water. Res., 35, 1518-1524, 2001.

Stanier, R.Y., Palleroni, N.J., Doudoroff, M.: The aerobic pseudomonads: a taxonomic study, J. Gen. Microbiol., 43, 159-271, 2010.

Va ïilingom, M., Amato, P., Sancelme, M., Laj, P., Leriche, M., Delort, A.M.: Contribution of microbial activity to carbon chemistry in clouds, Appl. Environ. Microbiol., 76, 23-29, 2010.

Va ïilingom, M., Attard, E., Gaiani, N., Sancelme, M., Deguillaume, L., Flossmann, A.I., et al.: Long-term features of cloud microbiology at the puy de Dôme (France), Atmos. Environ., 56, 88-100, 2012.

Va ïilingom, M., Deguillaume, L., Vinatier, V., Sancelme, M., Amato, P., Chaumerliac, N., et al.: Potential impact of microbial activity on the oxidant capacity and organic carbon budget in clouds, Proc. Natl. Acad. Sci. USA., 110, 559-564, 2013.

Wang, X., Chen, J., Sun, J., Li, W., Yang, L., Wen, L., et al.: Severe haze episodes and seriously polluted fog water in Ji'nan, China, Sci. Total. Environ., 493, 133-137, 2014.

Wolf, A., Fritze, A., Hagemann, M., Berg, G.: *Stenotrophomonas rhizophila* sp. nov., a novel plant-associated bacterium with antifungal properties, Int. J. Syst. Evol. Micr., 52, 1937-1944, 2002.

Womack, A.M., Bohannan, B.J., Green, J.L.: Biodiversity and biogeography of the atmosphere, Philos. Tran. R. Soc. Lond. B. Biol. Sci., 365, 3645-3653, 10.1098/rstb.2010.0283, 2010.

Wu, J., Peters, B.A., Dominianni, C., Zhang, Y., Pei, Z., Yang, L., et al.: Cigarette smoking and the oral microbiome in a large study of American adults, ISME. J., 10, 2435-2446, 10.1038/ismej.2016.37, 2016.

Xu, P., Yu, B., Li, F.L., Cai, X.F., Ma, C.Q.: Microbial degradation of sulfur, nitrogen and oxygen heterocycles, Trends. Microbiol, 14, 398-405, 2006.

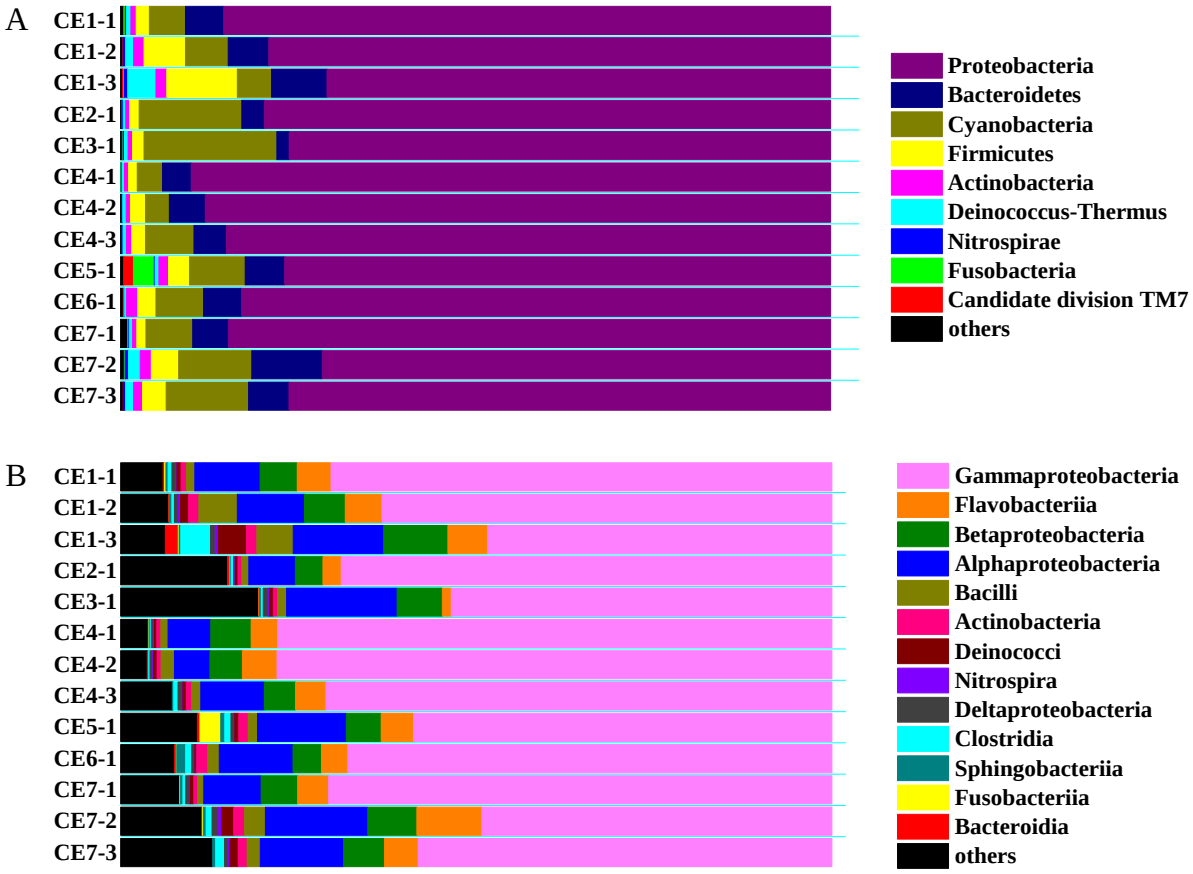


Figure 1 Bacterial community variation for cloud episodes at the phylum (A) and class (B) level. CE refers to the cloud episodes. Bar graphs for each sample represent the percentage of taxa assigned to each phylum with 80% bootstrap confidence. Taxonomic summary of the most abundant taxa (more than 1%) across all cloud samples are indicated in the figure.

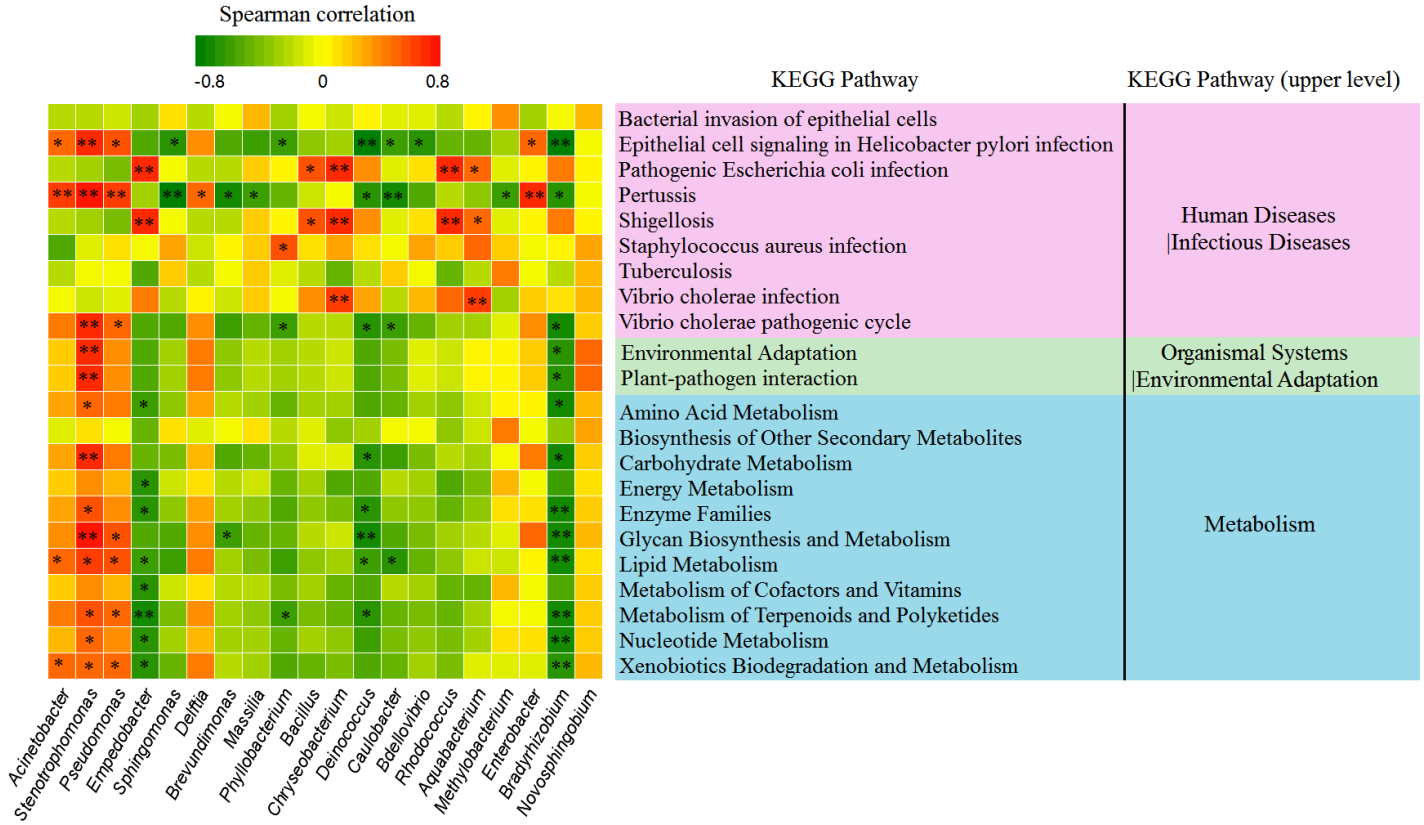


Figure 2 Bacterial taxa are related to KEGG functional pathways. Bacterial gene functions were predicted based on 16S rRNA gene sequences using the PICRUSt algorithm and annotated from KEGG databases. Spearman's correlation coefficients were calculated for each pairwise comparison of genus and KEGG pathway. Selected KEGG pathways related to metabolism and disease infection and predominant genera are included in the heatmap. Red color refers to the positive correlation, and green indicates a negative correlation. Correlation is significant at *P < 0.05, **P < 0.01.

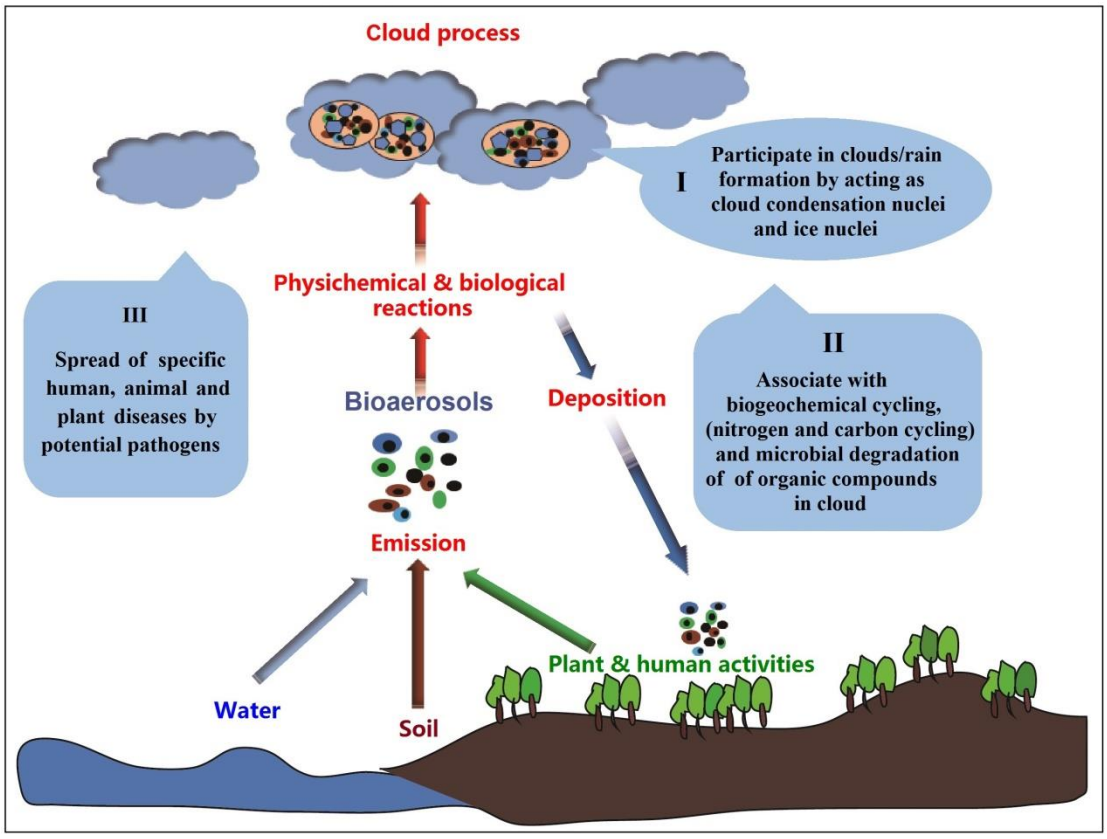


Figure 3 Schematic representation of bioaerosols life cycle and potential influence on atmosphere, ecosystem and human health, modified from Poeschl (Poeschl, 2006). The predominant bacteria species with potential functions are indicated in the figure.

Bioaerosols emissions and resuspension from various terrestrial environments, e.g., soil, water, plants, animals or human beings, may include pathogenic or functional species. These bacteria can be attached to particles or incorporated into water droplets of clouds/fog. Certain species can serve as biogenic nuclei for Cloud Condensation Nuclei (CCN) and Ice Nuclei (IN), which induce rain formation, precipitation, and wet deposition of gases and particles. For the potential pathogens and functional bacteria, during cloud process, they can be deposited back to land via deposition and possibly induce infections to human health and impose effect on the diversity and function of aquatic and terrestrial ecosystems.

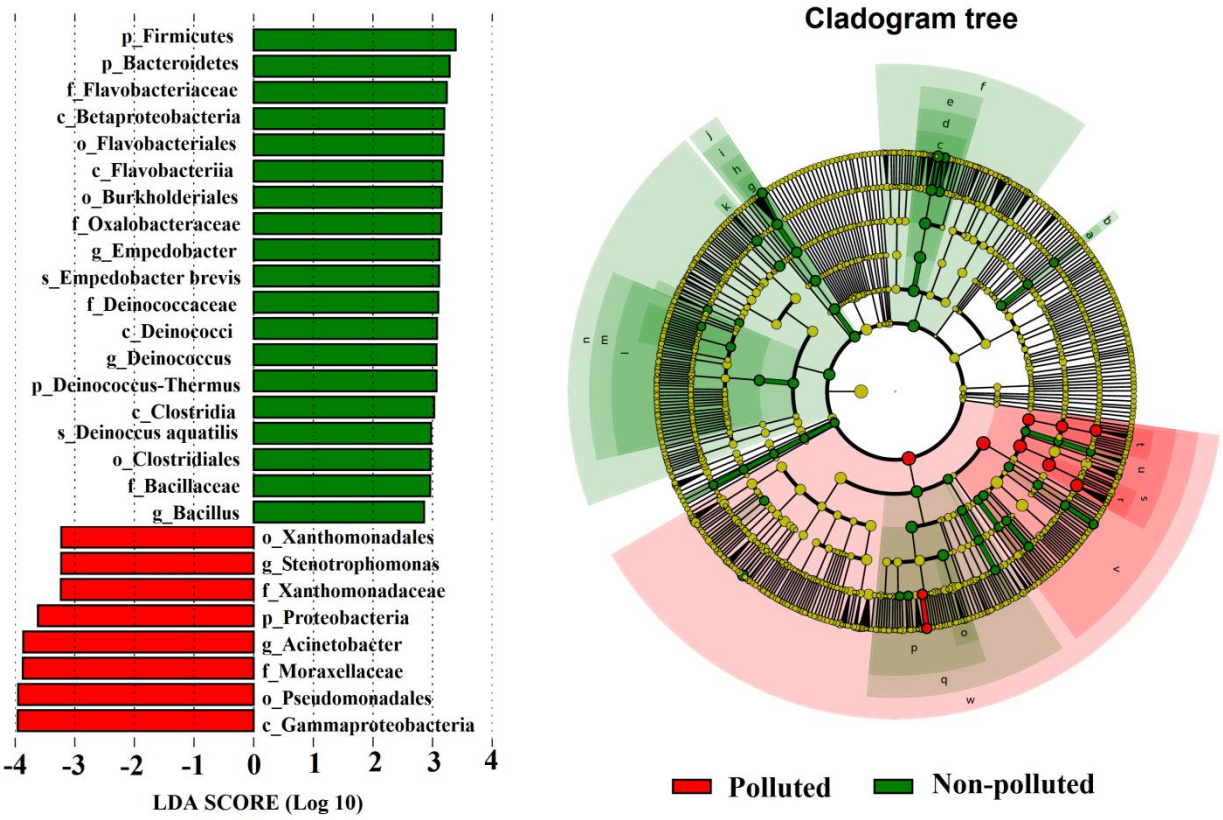


Figure 4 Distinct bacterial taxa between polluted and non-polluted cloud episodes identified by linear discriminant analysis coupled with effect size (LEfSe). The LDA effect sizes (left) were calculated using the default parameters. The taxonomic cladogram (right) was visualized with LDA values higher than 3.5 comparing all bacterial taxa. The significantly distinct taxon nodes are colored in red and green for polluted and non-polluted cloud episodes, respectively. Bacterial taxa with nonsignificant differences are indicated with yellow circles and circle diameter is proportional to relative abundance. The abbreviation in the cladogram tree: a: g_Rhodococcus, b: f_Nocardiaceae, c: f_Flavobacteriaceae, d: o_Flavobacteriales, e: c_Flavobacteriia, f: p_Bacteroidetes, g: f_Deinococcaceae, h: o_Deinococcales, i: c_Deinococci, j: p_Deinococcus-Thermus, k: f_Bacillaceae, l: o_Clostridiales, m: c_Clostridia, n: p_Firmicutes, o: f_Oxalobacteraceae, p: o_Burkholderiales, q: c_Betaproteobacteria, r: f_Moraxellaceae, s: o_Pseudomonadales, t: f_Xanthomonadaceae, u: o_Xanthomonadales, v: c_Gammaproteobacteria, w: p_Proteobacteria.

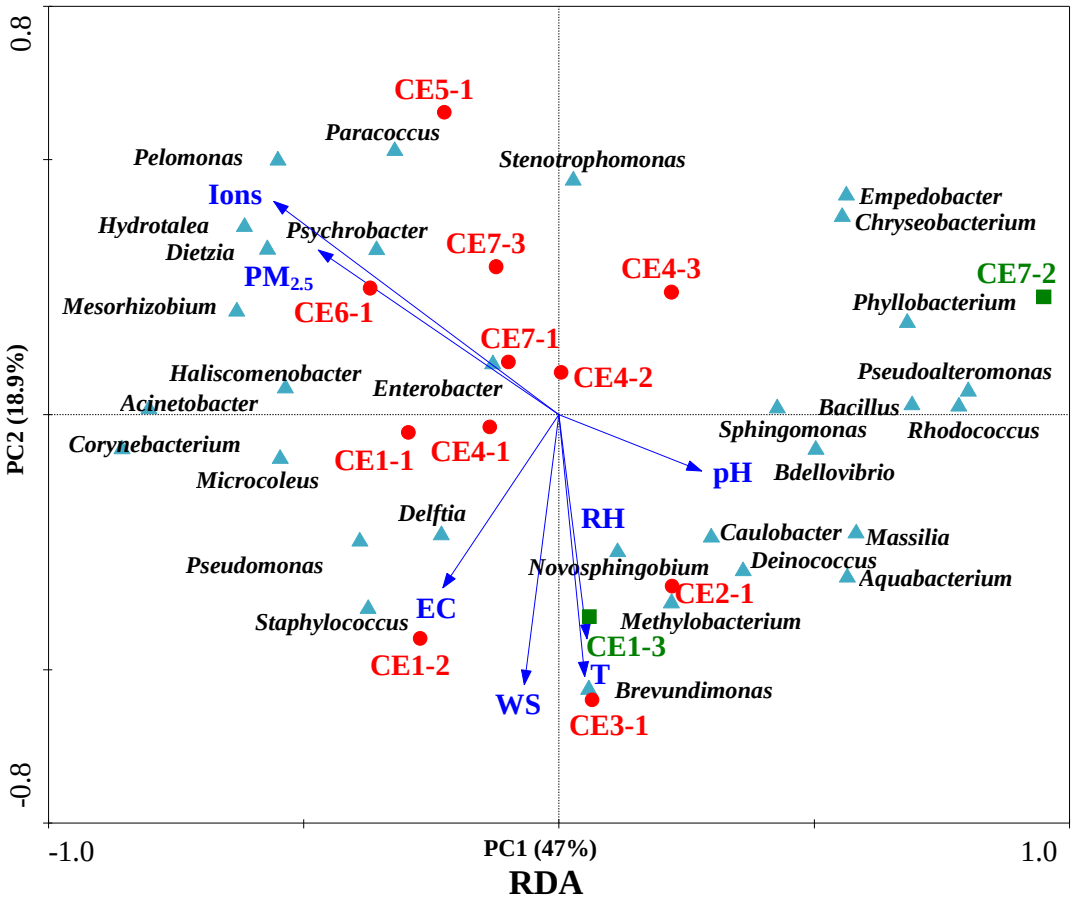


Figure 5 Biplot of the environmental variables and genus-level community structure using a redundancy analysis (RDA) model, describing the variation in bacterial community explained by environmental variable. CE refers to cloud episodes. Polluted episodes are indicated in red circle, and non-polluted episodes are green squares. Species data are listed in Table S2. The selected environmental variables are significant ($P < 0.05$) using Monte Carlo permutation testing. Species are labeled with triangle, the closer of which indicates existence in similar environment. Environmental variables are showed by arrows; the relative length is positive correlation with the importance in influencing bacterial community structure. The angle between the arrow and the ordination axis suggest the variable response respect to the RDA gradient. The two axes explain 65.9% of the variability. For bacteria, major ions in cloud water seem to be the most important environmental variable shaping the community structure.

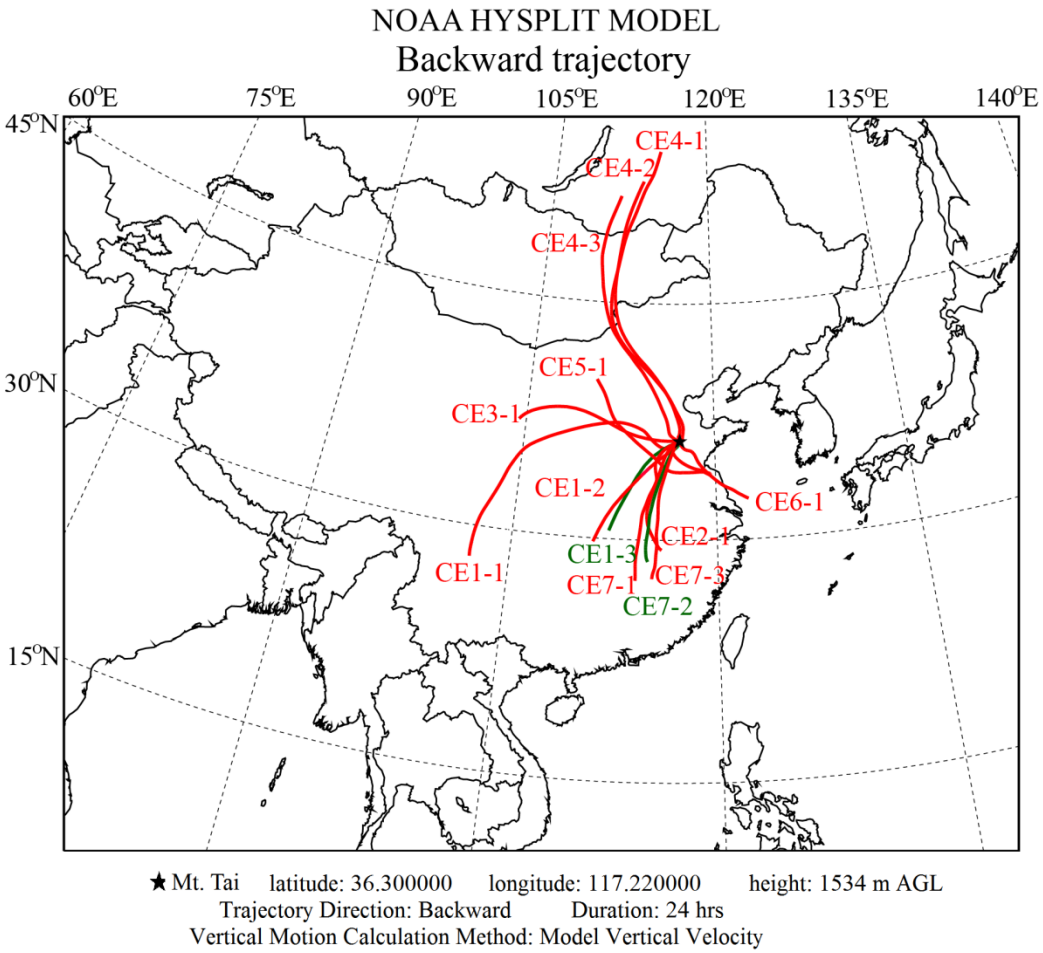


Figure 6 Air mass transport pathways for the cloud episodes using the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model. 24-hour backward trajectories were calculated for air parcels arriving at the summit of Mt. Tai (36°18′ N, 117°13′ E, and 1534 m a.s.l). CE refers to cloud episodes. The polluted episodes are indicated in red lines, and green lines are non-polluted episodes.

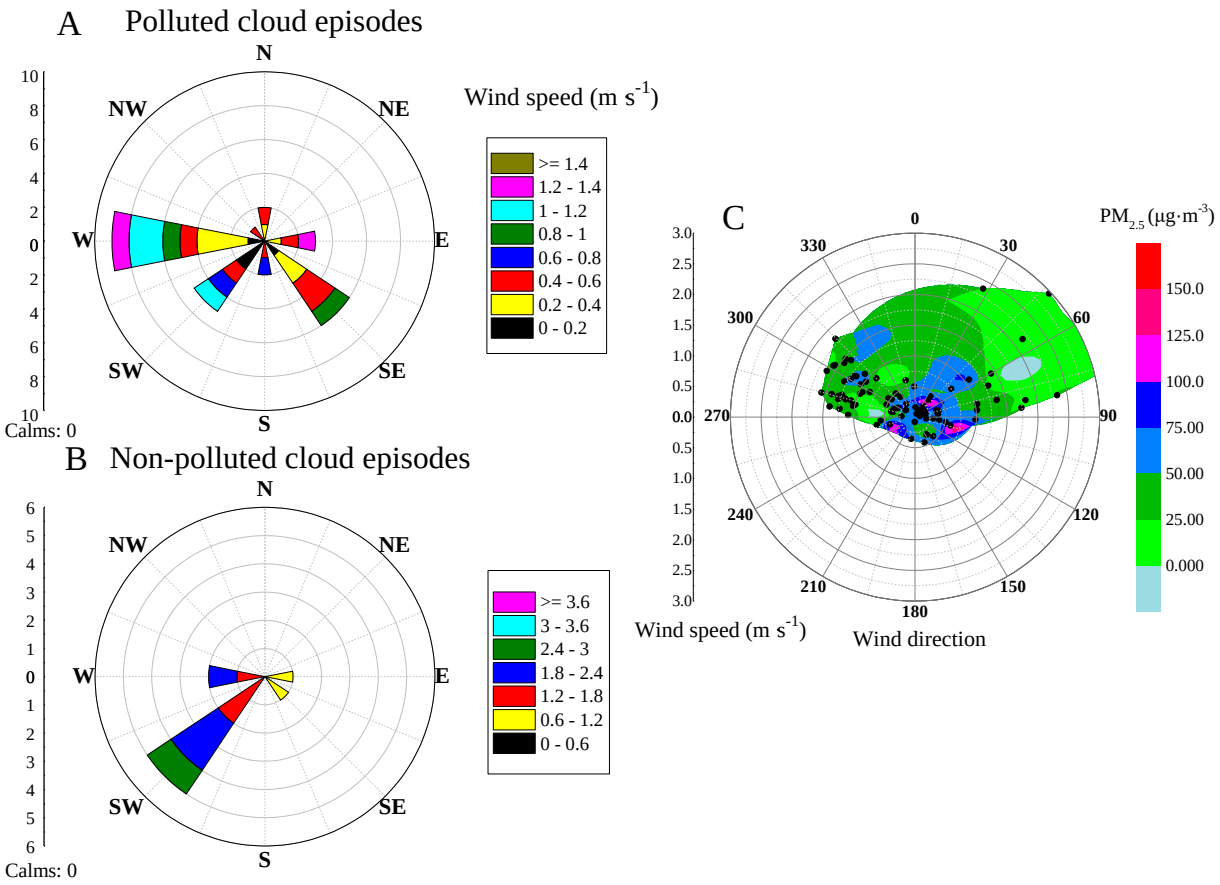


Figure 7 Wind Rose Diagram to quantitative analysis of wind speed and wind direction during sampling time between polluted (A) and non-polluted cloud episodes (B). The frequency of winds is indicated by wind direction. Wind speed range is labeled with color bands. Wind direction with the greatest frequency is shown with the direction of the longest spoke. Figure 7 (C) indicates distribution of wind speed during the whole sampling time (from 24 July to 23 August 2014) and correlation with $PM_{2.5}$ concentration. As shown in the figure, $PM_{2.5}$ concentration was high under lower wind speed, whereas $PM_{2.5}$ was lower when wind speed was high.

Table 1 Description of cloud episodes at Mt. Tai, China

Data	Samples	Start time (BJT)	Stop time (BJT)	Duration (h)	$PM_{2.5}$ [a] ($\mu g \cdot m^{-3}$)	LWC ($g\ m^{-3}$)	pH	EC ($\mu S \cdot cm^{-1}$)	OC ($mg\ L^{-1}$)
24 Jul 2014	CE1-1	8:50	15:30	6:40	105.07	0.21	4.03	583	ND
	CE1-2	15:30	17:30	2:00	22.35	0.23	4.32	219.2	ND
	CE1-3	17:30	22:51	5:21	14.66	0.24	5.74	104.4	ND
5 Aug 2014	CE2-1	6:45	9:17	2:32	30.36	0.22	5.80	275.7	ND
5 Aug - 6 Aug 2014	CE3-1	19:05	4:01	8:56	42.25	0.10	5.10	501	ND
14 Aug - 15 Aug 2014	CE4-1	22:41	0:44	2:03	42.69	0.02	6.36	170.4	BDL
	CE4-2	0:44	5:06	4:22	47.98	0.03	5.34	86.34	0.04
	CE4-3	5:06	6:03	0:57	36.88	0.02	4.89	64.95	BDL
17 Aug 2014	CE5-1	10:10	11:18	1:08	63.18	0.39	5.20	120.5	0.11
17 Aug - 18 Aug 2014	CE6-1	22:18	1:25	3:07	54.33	0.10	3.80	321.8	0.02
23 Aug 2014	CE7-1	2:30	4:38	2:08	30.45	0.20	4.38	356.2	0.03
	CE7-2	4:38	6:21	1:43	23.39	0.22	5.01	207.5	0.15
	CE7-3	6:21	9:20	2:59	41.60	0.21	5.74	187.6	0.21

CE refers to cloud episode. BJT refers to Beijing Time, which equals UTC + 8. LWC refers to the cloud liquid water content. EC refers to the electric conductivity. OC refers to the organic carbon in cloud water. ND means not detected due to instrument failure. BDL means below detection limitation.

Table 2 Summary of bacterial diversity and richness of cloud water

Sample ID	Reads	OTUs	Ace	Chao1	Coverage	Shannon	Simpson
Polluted cloud episodes							
CE1-1	18213	975	1835	1491	0.9761	3.9418	0.0646
CE1-2	18702	1184	1841	1730	0.9719	4.1919	0.0620
CE2-1	19914	1125	1756	1684	0.9744	3.9582	0.0630
CE3-2	18199	1022	2082	1582	0.9734	3.9749	0.0647
CE4-1	18350	941	1828	1461	0.9762	3.6041	0.0953
CE4-2	17707	967	1522	1427	0.9752	3.6748	0.0902
CE4-3	17397	981	2091	1611	0.9725	3.8074	0.0832
CE5-1	16384	1132	1814	1790	0.9676	4.3173	0.0546
CE6-1	16896	1186	1997	1872	0.9657	4.1268	0.0666
CE7-1	16350	1103	2501	1795	0.965	3.9040	0.0810
CE7-3	18122	1258	1958	1999	0.9686	4.3776	0.0531
Non-polluted cloud episodes							
CE1-3	17662	1173	1689	1687	0.9732	4.7067	0.0327
CE7-2	18252	1150	1732	1673	0.9729	4.3709	0.0426
Aerosol (Katra et al., 2014)	4020	1412		2142	0.8300		
Bioaerosol (Madsen et al., 2015)						2.64-3.05	0.816-0.922
Rain water in July (Cho & Jang, 2014)	3055	1542	13083	6387			
$PM_{2.5}$ in summer (Franzetti et al., 2011)		2222		4036			
TSP annual (Bertolini et al., 2013)	271587	765-26,187		107		2.40	

The diversity indexes including OTUs, ace, Chao1, coverage, shannon, simpson were defined at 97% sequence similarity. CE refers to cloud episodes. TSP refers to the total suspended particulate matter.

Table 3 The identified bacterial species in cloud water samples correlation with the potential ecological function

Genus	Identified species	Habitats	Ecological role	Reference
Acinetobacter GP	*A. schindleri*	soil/water	CNN or IN; Opportunistic pathogens	(Mortazavi et al., 2008; Nemec et al., 2001)
Bacillus FR	*B. anthracis*	soil/water/air	CNN or IN; Opportunistic pathogens	(Makino & Cheun, 2003; Mortazavi et al., 2008)
Brevundimonas BP	*B. diminuta*	soil/water	CNN	(Bauer et al., 2003; Han & Andrade, 2005)
	B. vesicularis	soil/water	Opportunistic pathogens	(Gilad et al., 2009)
Caulobacter AP	*Caulobacter. sp.*	water	Metabolism/Biodegradation	(Nakamura et al., 2007)
Chryseobacterium BA	*C. aquaticum*	soil/water	Protect and promote plants growth	(Gandhi et al., 2009)
	C. jejuense	soil/water		(Ben Abdeljalil & Vallance, 2016)
Clostridium FR	*C. tertium*	soil/gut	Opportunistic pathogens	(Miller et al., 2001)
Comamonas BP	*C. testosteroni*	soil/water	Metabolism/Biodegradation	(Goyal & Zylstra, 1996)
Cyanobacterium CY	*Cyanobacterium sp.*	soil/water	Carbon and nitrogen fixing	(Jha et al., 2004)
Deinococcus DT	*D. aquatilis*	soil/water	Extremophiles, radiation-resistant	(Kämpfer et al., 2009)
Delftia BP	*D. tsuruhatensis*	soil/water	Metabolism/Biodegradation	(Geng et al., 2009)
Empedobacter BA	*E. brevis*	soil/water/plant	Opportunistic pathogens	(Bottone et al., 1992)
Methylobacterium AP	*M. aquaticum*	water	Methylotrophic, carbon fixing	(Gallego et al., 2005)
	M. adhaesivum	soil/water		(Gallego et al., 2006)
Moraxella GP	*M. osloensis*	soil/animal	Opportunistic pathogens	(Banks et al., 2007)
Novosphingobium AP	*N. aromaticivorans*	soil/water	Metabolism/Biodegradation	(Bell & Wong, 2007)
Staphylococcus GP	*S. equorum*	soil/water/clinic	Opportunistic pathogens	(Novákov á et al., 2006)
Phyllobacterium AP	*P. myrsinacearum*	soil/plant	Rhizosphere bacteria, nitrogen fixation	(Gonzalezbashan et al., 2000)
Pseudomonas GP	*P. psychrotolerans*	soil/water	Extremophiles, psychrotolerant	(Hauser et al., 2004)
	P. geniculate	soil/water/plant	Metabolism/Biodegradation	(Gopalakrishnan et al., 2015; Liu et al., 2014)
	P. putida	water/soil	Protect and promote plants growth	(Meziane et al., 2005; Reardon et al., 2000)

Table 3 (Continued)

Genus	Identified species	Habitats	Ecological role	Reference
			CNN or IN	(Amato et al., 2015; Joly et al., 2013)
Rhodococcus AC	*R. ruber*	soil/water	Metabolism/Biodegradation	(Bock et al., 1996)
Sphingomonas AP	*S. faeni*	soil/water	CNN or IN; psychrotolerant	(Ponder et al., 2005)
	S. kaistensis	soil/water	Metabolism/Biodegradation	(Busse et al., 2003)
	S. leidyi	soil/water		(Glaeser & Kämpfer, 2014)
Stenotrophomonas GP	*S. rhizophila*	soil/water/plant	CNN or IN; Rhizosphere bacteria	(Mortazavi et al., 2008; Wolf et al., 2002)

CNN and IN refers to the bacteria participating in the formation of clouds or rain by acting as cloud condensation nuclei (CNN) and ice nuclei (IN).

Abbreviates are as followed: AP, Alphaproteobacteria; BP, Betaproteobacteria; GP, Gammaproteobacteria; AC, Actinobacteria; BA, Bacteroidetes; CY, Cyanobacteria; DT, Deinococcus-Thermus; FR, Firmicutes.

Biodegradation refers to the bacteria associated with the biodegradation of organic compounds, even toxic pollutants, e.g., aromatic compounds.

The initially submitted manuscript

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

Min Wei [1], Caihong Xu[1], Jianmin Chen[1,2,*], Chao Zhu[1], Jiarong Li[1], Ganglin Lv[1]

[1] Environment Research Institute, School of Environmental Science and Engineering, Shandong University, Ji'nan 250100, China

[2] Shanghai Key Laboratory of Atmospheric Particle Pollution and Prevention (LAP), Fudan Tyndall Centre, Department of Environmental Science & Engineering, Fudan University, Shanghai 200433, China

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

Abstract:

Bacteria, widely distributed in atmospheric bioaerosols, are indispensable component in fog water system and play an important role in atmospheric hydrological cycle. However, little is known about the bacterial community dynamics and ecological function, especially under the increasing serious air pollution events in North China Plain. Here we have a comprehensive characterization of bacterial community structure, variation and environmental influence about fog water collected at Mt. Tai under polluted and non-polluted fog episodes from 24 Jul to 23 Aug 2014. Using the Miseq 16S rRNA gene sequencing, the facts that fog water harbored a highly diverse bacterial community and the predominant phyla of Proteobacteria, Bacteroidetes, Cyanobacteria and Firmicutes were investigated. The abundant genera *Acinetobacter*, *Stenotrophomonas*, *Pseudomonas*, and *Empedobacter* originated from a wide range of

habitat included opportunistic pathogenic and functional species, suggesting the bacterial ecological and healthy importance in fog water should be concerned. Clustering analysis including hierarchical cluster (Hcluster) and principal coordinate analysis (PCoA) indicated a significant disparity between polluted and non-polluted samples. Potential pathogens were significant group in the polluted samples, whereas a more diverse ecological function group of bacteria were identified in the non-polluted samples using linear discriminant analysis effect size (LefSe). Community structure discrepant performed by redundancy analysis (RDA) indicated $PM_{2.5}$ have negative impact on bacteria, playing vital role in shaping microbial community structure. $PM_{2.5}$ was possibly associated with different origins and pathways of air mass using source tracking by the backward trajectory and wind analysis, mainly related to the long-term transport combing with local regional emission processes. This work furthered our understanding of bacterial ecological characteristics in the atmospheric aqueous phase, highlighted the potential influence of environmental variables on bacterial community over fog process, which will provide fundamental acquaintance of bacterial community response in fog water under increasing pollution stress.

Key words： fog water, bacterial diversity, community disparity, $PM_{2.5}$

1. Introduction

Fog is the near-surface cloud and aerosol system composed of tiny droplets suspended in the atmosphere. In the atmosphere, numerous pollutants could be dissolved or suspended in fog, which may induce complex effects on environment security and human health. Over the past decades, studies on fog/cloud water are mainly focused on the physicochemical properties (Aikawa et al., 2001; Boris et al., 2015; Fernández-González et al., 2014). Recently, with the in-depth understanding of the characteristics of fog, bioaerosols in fog have been the upcoming focus.

Studies have showed that living microorganisms, including bacteria, fungi and yeasts, are present in fog or clouds (Burrows et al., 2009). As the first study on biological particles in fog water, Fuzzi et al (1997) suggest the bacterial replication in foggy days. Afterwards, with the development of detection techniques, microorganisms in fog/cloud water are more systematically studied (Amato et al., 2007c; Delort et al., 2010; Vaïtilingom et al., 2012). Combined with the field investigations and lab experiments, diverse bacterial communities are identified, and the bacterial metabolically active in fog/cloud water are also demonstrated. In atmospheric aqueous phase, microorganisms can act as cloud condensation nuclei and ice nucleation, which have potential impact on precipitation processes (Amato et al., 2015; Mortazavi et al., 2015). Moreover, microorganisms in fog/cloud water are available to metabolize organic carbon compounds and influence photochemical chemical reactions (Vaïtilingom et al., 2013), involve in the nitrogen cycling (mineralization and nitrification) (Hill et al., 2007), degrade organic acids (formate, acetate, lactate, succinate) and associate with carbon recycling (Amato et al., 2007a; Vaïtilingom et al., 2010), and therefore participate in a series of complex and diverse biochemical metabolic activities.

A fog occurrence is a complex process, in contaminated area, fog typically contains numerous pollutants, e.g., sulfate and nitrate ions, organic carbon compounds, and bacteria(Badarinath et al., 2007; Després et al., 2012; Fernández-González et al., 2014; Mohan & Payra, 2009). Emissions and resuspension of bacteria by wind erosion or splashing water from various terrestrial environments into the atmosphere recruit diverse airborne bacteria, which possibly involve opportunistic and functional bacteria. During fog process, these bacteria attached to particles or incorporated in fog water

will be deposited back to the land via dry or wet deposition processes, which may induce human risks through microbial pathogens dispersion and potential effect on the diversity and function of aquatic and terrestrial ecosystems. Therefore, to evaluate the potential ecological functional bacteria in fog water is urgent, especially for the polluted fog episodes.

It is noteworthy that airborne bacterial communities are closely related to environmental characteristics (Gao et al., 2016), and meteorological factors are often correlated with the observed bacterial community structure (Dong et al., 2016). For instance, studies about the relationships between ambient inhalable airborne and environmental parameters suggest temperature, relative humidity, PM_{10}, $PM_{2.5}$ and particle size have significant impact on the composition and dynamic of microbial communities (Adhikari et al., 2006; Bowers et al., 2013). However, due to the paucity of detailed and comprehensive studies of atmospheric bacterial composition, the understanding of the dynamic of bacterial community remains incomplete, particularly in the North China Plain. The North China Plain is the most important agricultural and economic region in China, which has been suffering serious air pollution events in recent years, e.g., the severe fog and haze pollution in Beijing during January 2013 (Huang et al., 2014). During a polluted fog process, how bacterial community varied and which environmental factor play decisive role in shaping bacterial community structure are still unclear.

In the present work, typical fog episodes under polluted and non-polluted weather were collected in the summit of Mt. Tai in North China Plain. To understand the dynamic of bacterial community, the Miseq 16S rRNA gene sequencing was performed, and analysis of similarities (ANOSIM) and linear discriminant analysis effect size (LEfSe) were executed to clarify the discrepant bacterial taxa. Moreover, RDA analysis was applied to identify the pivotal environmental factor influencing bacterial community. Air mass back trajectory and wind direction and speed analysis were selected to definitude the most likely source and transmission paths of pollutants and bacteria.

2. Material and methods

2.1 Sample collection

Fog samples were collected using the Caltech Active Strand Cloud water Collector (CASCC2) with a droplet size cut of 3.5 μm at the summit of Mt. Tai (36°18′ N, 117°13′ E, and 1534 m a.s.l) (Guo et al., 2012). The flow rate was 24.5 $m^3 min^{-1}$ and fog water was collected on the strings flows down to Teflon bottles. The collected samples were stored at 4°C until analysis.

In fog water, the pH and conductivity was detected with a Multi pH/COND/TEMP 6350 hand held Meter immediately after sampling. Hourly data, e.g., meteorological parameters, and $PM_{2.5}$ were measured to evaluate the air quality during fog episodes. The meteorological parameters including atmospheric visibility, temperature, relative humidity, wind direction, wind speed were measured with an automatic meteorological station (PC-4, JZYG, China) *in situ*. The mass concentration of $PM_{2.5}$ was measured using a Model 5030 SHARP monitor (SHARP 5030, Thermo Fisher Scientific, Massachusetts, USA). To determine the most likely source region for air mass of fog episodes, the 24-h back trajectory analysis was performed using the Hybrid Single-Particle Lagrangian Integrated Trajectories (HYSPLIT) model (http://ready.arl.noaa.gov/HYSPLIT.php). Moreover, the wind rose diagram of study area (origin, version 9.0, Origin Lab Corporation, Northampton, MA)) during fog process were utilized to clarify the predominant wind direction and wind speed.

2.2 DNA Extraction and PCR Amplification

Genomic DNA was extracted in triplicate with the FastDNA spin kit for soil (MP Biomedicals, Solon, OH, USA) according to the manufacturer's directions. The concentration of DNA was determined spectrophotometrically (Nano-Drop 2000, Thermo, Wilmington, Delaware, USA).

The designed primer sets with the V3-V4 region of 16S rRNA gene (338F-806R) (Masoud et al., 2011), adapter and barcodes were selected in the illumina Miseq sequencing. For each sample, a 25-μL PCR mix was prepared and contained 10 μL of 5x Buffer, 1μL of dNTP (10mM), 1 U Phusion High-Fidelity DNA polymerase, 20 ng

of template DNA, 1 μL of each 10 μM modified primer, with double-distilled water until 25 μL. PCR was performed at 94°C for 2 min; 25 cycles of 94 ℃ for 30 s, 56 ℃ for 30 s and 72 ℃ for 30 s; 72 ℃ for 5 min; and hold at 10 ℃.

The PCR products were separated by 2% agarose gel electrophoresis and purified with the nucleic acid purification Kit (AxyPrepDNA, Axygen, USA). The purified PCR products were quantified using a Qubit 3.0 fluorometer (Invitrogen, Carlsbad, CA) and then mixed to the equal concentration. For each sample, 4 μL of 10 nM pooled DNA was denatured with 1 μL of 0.2 N NaOH at room temperature. Finally, the Illumina paired-end sequencing was performed on a MiSeq platform (Illumina, Inc., San Diego, CA). After sequencing, two FASTQ files (read 1 and read 2) for each sample were generated on the MiSeq reporter software automatically. Raw 16S rRNA gene sequences are available at the Sequence Read Archive (SRA) under accession numbers SRX1904235.

2.3 Illumina high-throughput sequencing and analyzing

Raw sequences were processed and analyzed using the QIIME package (Kuczynski et al., 2011). The PE reads were firstly merged with overlap greater than 10 bp. Then, the adapter, barcodes and primers were removed from the merged sequences. Subsequently, the trimmed sequences with length shorter than 200 bp, quality score lower than 25, homologous longer than 8 bp, contained ambiguous characters were screened. Finally, chimeric sequences were identified using the Usearch61 algorithm and removed from the dataset. The optimized sequences were clustered into OTUs at the threshold of 97% similarity with the usearch61 algorithm. Single OTUs were removed and taxonomy was assigned to each OTU using the Ribosomal Database Project (RDP) classifier in QIIME, with a minimum confidence cutoff of 0.8 against the Silva reference database (silva 119,http://www.arb-silva.de/) to the genus level.

Alpha diversity was assessed by examining the rarefaction curves, shannon-wiener curves and rank-abundance curves calculated with Mothur (v.1.34.0; http://www.mothur.org) (Schloss et al., 2009) and visualized in R project (v.3.1.3; https://www.r-project.org/). Community richness estimators including the observed OTUs, nonparametric Chao1, ACE, and community diversity estimators including Shannon and Simpson indexes were also calculated with Mothur. Moreover, the Good's coverage was used to evaluate the sequencing depth.

Differences between polluted and non-polluted samples were tested by ANOSIM (Clarke, 1993). The ANOSIM R statistic is calculated on the basis of difference in mean ranks between and within groups. Linear discriminant analysis effect size (LEfSe, http://www.huttenhower.sph.harvard.edu/galaxy/) was applied to identify differentially abundant bacterial taxa associated with the polluted and non-polluted fog episodes at genus or higher taxonomy levels (Segata et al., 2011).

2.4 Interaction between bacterial community structure and environmental variables

To determine the relationship between bacterial community structure and environmental variables, a detrended correspondence analysis (DCA) was first performed to estimate the length of the gradient. The resulting length (0.99) indicates a linear model was appropriate, hence RDA was subsequently performed. RDA was elaborated with the predominant bacteria data matrix and the environmental data matrix (Anderson & Willis, 2008). Interset correlations of this analysis were used to determine which environmental variables were the most important in determining the community structure. The cumulative fit per species as fraction of variance of species was performed to determine the importance of a species for the ordination space and which species were most associated with environmental factors. Analysis was performed in Canonical Community Ordination for windows (Canoco, v 4.5).

3. Results and discussions

3.1 Microbial community in fog water

Seven fog episodes from 24 July to 23 August 2014 were observed. Detail information was summarized in Table 1. Fog episodes can be classified as polluted and non-polluted according to the average $PM_{2.5}$ mass concentration.

Information on the bacterial community of fog water has been very scarce, our study provided comprehensive investigation of bacterial community under both polluted and non-polluted fog episodes. From the 13 samples collected during 7 fog episodes, a total of 232148 high quality sequences were obtained after quality filtering and OTUs ranged from 975 to 1258 (Table 2). This value was similar with the previous sequence-based survey of atmospheric bacteria (OTUs, 1214) (Katra et al., 2014).

Identification of OTUs at different taxonomic levels yielded 359 species, 411 genera, 152 families, 70 orders, 38 classes and 26 phyla.

Rarefaction curves of observed OTUs continued to rise with increasing numbers of sequences (Figure S1), 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 (1671) was similar to the non-polluted samples (1696). Diversity estimators Shannon and Simpson indexes fluctuated between polluted and non-polluted samples. Bacterial diversity was higher in non-polluted samples (polluted, 3.94; non-polluted, 4.42).

Across all samples collected from the 7 fog episodes, Proteobacteria was the dominant phylum, followed by Bacteroidetes, Cyanobacteria, Firmicutes, Deinococcus-Thermus, Actinobacteria and Nitrospirae (Figure 1). The bacterial community structure is similar to few other studies explored the bacterial diversity in cloud/fog samples, the aforementioned phyla contained a series of genera participate in the atmospheric hydrological cycle (Amato et al., 2007b; Delort et al., 2010). They are predominant taxa in clouds at a high elevation determined by sanger sequencing and tagged pyrosequencing (Bowers et al., 2009), and which are also the typical culturable heterotrophic bacteria from clouds broadly distributed in aquatic and terrestrial habitats (Amato et al., 2005; Kourtev et al., 2011). In the present study, Figure 2 shows the dominant genera collected during fog process. For the 7 fog episodes, the predominant genera from Proteobacteria were similar, including *Acinetobacter*, *Stenotrophomonas*, *Pseudomonas*, *Sphingomonas*, *Massilia*, *Delftia*, *Brevundimonas*. These bacteria contained a metabolically diverse group found in a wide range of water/soil habitats. For instance, *Acinetobacter*, the most abundant genera widely distributed in land or ocean, are contribute to the biodegradation, leaching and removal of several organic and inorganic man-made hazardous wastes (Abdelelhaleem, 2003). *Stenotrophomonas* and *Pseudomonas*, which are well-known for the striking capability to utilize numerous carbon sources, have been widely

utilized in the degradation and transformation of complex organic compounds in a wide range of habitats (Boonchan et al., 1998; Stanier et al., 2010). *Sphingomonas*, has reported the ability to survive in low concentrations of nutrients, metabolize a series of carbon compounds, events toxic compounds (Xu et al., 2006). Similar to *Sphingomonas*, members of *Brevundimonas* are well known for their ability to withstand extreme harsh environment (Kopcakova et al., 2014). *Massilia*, isolated from air samples, could participate in the biodegradation and transport of Phenanthrene (Gu et al., 2016). *Empedobacter* from Bacteroidetes are widely distributed in water habitats, since the human clinical origins, *Empedobacter* are ranked as potential pathogens (Hugo et al., 2005). *Bacillus* included in the phylum Firmicutes commonly found in soil and water, are also found in air samples (Suominen et al., 2001). Similar to *Pseudomonads*, some strains of *Bacillus* could produce biosurfactants that can act as cloud condensation nuclei (Delort et al., 2010). Moreover, members of *Deinococcus* from Deinococcus-Thermus are well known for their ability to withstand extreme radiation conditions that could potentially adapt to the cloud environment (Mattimore & Battista, 1996). The identification of bacteria adapt to specific environments in fog/cloud water (low temperature, harsh nutrition and high radiation environment) with the potential role in the nucleation, metabolism of organic pollutants, demonstrated the potential importance of participation and influence the atmospheric biochemistry cycle.

3.2 Implications in human health and ecosystem

Bioaerosols have been complex assemblages of airborne and exogenic microorganisms, many of which likely emissions and resuspension from various terrestrial environments, e.g., soil, water, plants, animals or human beings. In the atmosphere, bacteria including pathogenic or beneficial species can be attached to particles or incorporated into water droplets of clouds/fog. During fog process, they can be deposited back to land via deposition and possibly induce infections to human health and impose effect on the diversity and function of aquatic and terrestrial ecosystems (Figure 3).

In the present study, the presence of potential pathogen sequences indicated occasional distribution and dispersion of pathogens in fog water. The levels of opportunistic pathogens found in polluted fog episodes are comparable to

non-polluted samples (Table 3). The identified opportunistic pathogens, e.g., *Empedobacter brevis*, can be easily isolated from clinical resources, which may be associated with eye infections (Bottone et al., 1992). Occurrence of *Staphylococcus equorum* in fog water can be expected since *Staphylococcus* are frequently isolated from airborne samples (Seo et al., 2008). As important pathogens, they can reside on the skin and mucous membranes of humans and induce server infections (Nováková et al., 2006). Similarly, the *Brevundimonas vesicularis* and *Brevundimonas diminuta* can induce virulent infections, often associated with nervous system or bacteraemia (Gilad et al., 2009; Han & Andrade, 2005). Besides that, the *Acinetobacter schindleri* and *Moraxella osloensis* are associated with skin and wound infections, bacteremia and pneumonia (Banks et al., 2007; Nemec et al., 2001). Previous studies has showed health risk-related bacteria in atmospheric samples, including rainwater (Cho & Jang, 2014), which can be part of atmospheric life cycle and dispersal pathway for some pathogenic bacteria. Possibly, greater survival of human pathogens may be supported in the atmosphere. Since the dispersion of these opportunistic pathogens via aerosol and fog droplets will cause infection of skin tissue and internal organs, pathogens in fog water need special attention.

The identified ecological function bacteria mainly participated in the biodegradation of organic compounds, such as *Rhodococcus ruber*, *Sphingomonas faeni*, *Delftia tsuruhatensis*, *Comamonas testosterone* (Table 3). Excessive studies have illustrated their capability of metabolism of hydrocarbon compounds, even toxic pollutants, e.g., aromatic compounds (Bock et al., 1996; Busse et al., 2003; Geng et al., 2009; Goyal & Zylstra, 1996). *Stenotrophomonas rhizophila* and *Phyllobacterium myrsinacearum* are two typical rhizospheric microorganisms. As plant-associated strains, S. rhizophila fulfill plant-protective roles and have been safely applied in biotechnology (Alavi et al., 2013). *Phyllobacterium myrsinacearum*, which is a predominant rhizospheric bacterium, its capability of azotification has made the utilization in plant growth promotion and biological control of soil-borne diseases (Gonzalezbashan et al., 2000). In addition to the potential impact on human health and ecosystem, there are extremophiles, e.g., *Deinococcus aquatili*, which is radiation-resistant and well adapted to the harsh atmospheric conditions (Kämpfer et al., 2009), *Pseudomonas psychrotolerans*, a psychrotolerant bacterium, could grow at 4^{o}C (Hauser et al., 2004). Overall, fog water seems to harbor highly diverse bacterial communities in ecosystem,

which may be the atmospheric mixing of diverse point sources origins in the rhizosphere, soil and water and possibly participate in the biodegradation of organic compounds in fog water.

3.3 Disparity between polluted and non-polluted fog episodes

Although the predominant bacteria are similar between polluted and non-polluted fog episodes, significant disparity within bacterial taxa are also identified. ANOSIM analysis suggest that the OTUs of the polluted samples were grouped into one large cluster, and were distantly related to the non-polluted clusters (ANOSIM comparison, $R = 0.579$, $P < 0.05$). Cluster analysis including PCoA and Hcluster indicated that the bacterial communities in polluted samples, regardless of the fog episodes, were highly similar (Figure S2). Cluster analyses based on the relative abundance of genera showed similar clustering patterns (Figure 4), and the polluted samples also shared high similarity in their bacterial community structure.

To find specialized bacterial groups within the polluted and non-polluted fog episodes, LEfSe is performed, which showed statistically significant differences. A total of 70 bacterial groups were distinct using the default logarithmic (LDA) value of 2. Cladograms show taxa with LDA values higher than 3.5 for clarity (Figure 5). Consequently, 8 and 19 differentially represent bacterial taxa in polluted and non-polluted fog episodes were detected.

In polluted fog episodes, most indicated bacteria were ranked as opportunistic pathogens, such as Proteobacteria, Gammaproteobacteria, Xanthomonadales, Xanthomonadaceae, Stenotrophomonas, Haliscomenobacter, and Saprospiraceae. Like Proteobacteria, the Gram-negative Gammaproteobacteria contains a series of ecologically and medically important bacteria, e.g., the pathogenic Enterobacteriaceae, Vibrionaceae and Pseudomonadaceae. The Xanthomonadales from this phylum has been reported to cause disease in plants (Saddler & Bradbury, 2005). Certain species from Stenotrophomonas (Gammaproteobacteria, Xanthomonadales) have been associated with a variety of infections in humans. The Saprospiraceae, a family within the phylum Bacteroidetes, include the genus *Haliscomenobacter*, are typical planktonic bacteria isolated from aquatic environments, such as marine, freshwater and activated sludge. The notable ability for the hydrolysis and utilization of complex carbon compounds has been illustrated (McIlroy & Nielsen, 2014).

In comparison, the majority of detected taxa in the non-polluted samples are from Bacteroidetes, Firmicutes, Alphaproteobacteria, Betaproteobacteria and Deinococcus-Thermus. An important biomarker from Bacteroidetes is Flavobacteriia, relative study has illustrated the marine sources for Flavobacteria, most of Flavobacteria sequences searched by NCBI blast are mainly from marine sources, i.e., algae, oysters and sea cucumbers (Cho & Jang, 2014). The genera *Empedobacter* was abundant across all samples, which are included in the family Flavobacteriaceae. As mentioned above, *Empedobacter* and *Staphylococcus* **(**Firmicutes) are potential pathogens and resistant to a wide range of antimicrobials (Hugo et al., 2005; Trilla & Miro, 1995). Phyllobacteriaceae (Alphaproteobacteria) are typical rhizobia. Similar with Phyllobacterium, members of Oxalobacteraceae (Betaproteobacteria) are rhizosphere microorganisms involved in the biological nitrogen fixation (Donn et al., 2015). The Burkholderiales (Betaproteobacteria) commonly found in water and soil are involved in the biodegradation of various forms of aromatic compounds (Pérez-Pantoja et al., 2012). Another important group of indicator bacteria was Deinococci from the phylum of Deinococcus-Thermus, which include many resistant species, as well as several thermophiles, could metabolize toxic materials, resistant to extreme radiation and survive in extremes of heat and cold (Griffiths & Gupta, 2007). Although bacteria with the potential function of nitrogen fixing (Phyllobacterium from Alphaproteobacteria) and degradation of organic compounds (Burkholderiales from Betaproteobacteria) were distinguished both in polluted and non-polluted episodes, a significant distinct was observed. By comparison, potential pathogens were significant groups in the polluted samples, whereas a diverse ecological function group of bacteria were identified in the non-polluted samples originated from a wide range of habitat. Ecologically meaningful distinguish of bacterial groups under polluted and non-polluted conditions is essential for understanding the structure and function of bacterial communities, and which provide a general understanding of the metabolism of bacteria in fog water.

3.4 Environmental factors shaping bacterial community structure

To clarity the vital environmental variable in shaping bacterial community structure, RDA was performed to discern the genus-level structure with the selected environmental factors (Figure 6). The first two *axes* explained 73.3% of the

accumulated variance in the species-environment relation. *Interset correlations* showed $PM_{2.5}$ was the most important environmental variable structuring the bacterial community (axis 1, -0.328); in turn, temperature registered the highest value for axis 2 (0.368) (Table S1). Cumulative fit indicated that the predominant genera affiliated with groups from *Acinetobacter*, *Empedobacter*, *Phyllobacterium*, *Aeromonas* and *Prevotella* displayed strong correlations with axis 1, *Streptococcus*, *Stenotrophomonas*, *Brevundimonas*, *Deinococcus* and *Pseudomonas* were the notable genera highly correlations with axis 2 (Table S2). As aforementioned in section 3.3, the above bacteria were metabolically diverse groups found in various habitats.

Of the atmospheric environmental characteristics measured, $PM_{2.5}$ was the best predictor of variability in diversity levels within the dominant genera and strongly correlated with represent bacterial genera. The composition of bacterial communities was highly variable under $PM_{2.5}$ mass concentration in this study, which was consistent with the previous study that $PM_{2.5}$ was important environmental factor shaping the variation of community composition (Cao et al., 2014). Moreover, statistical analysis, e.g., correlation or multiple linear regression, indicated that $PM_{2.5}$ exhibited a negative correlation with airborne bacteria during (Gandolfi et al., 2015; Gao et al., 2015) whereas in another study, spearman correlation analysis showed $PM_{2.5}$ exhibited a significant positive correlations with the airborne microbe concentration during hazy days (Dong et al., 2016). Previous study has suggest that nutrients are related to the distribution of bacteria in water habitats (Newton et al., 2011). Possibly, the inorganic and organic compounds in particulate matter ($PM_{2.5}$) can be available nutrients for microbial growth. However, aggregation of harmful substances such as heavy metals, polycyclic aromatic hydrocarbons would be toxic for bacteria under high $PM_{2.5}$ mass concentration. Since the $PM_{2.5}$'s two-sided influence on bacterial growth, bacterial community both under polluted and non-polluted samples were significantly correlated with $PM_{2.5}$ mass concentrations.

The identified taxa either from polluted or non-polluted samples were found in soil, water, plant or human skins. These bacterial groups are aerosolized and dispersed in the air, and partly from local regional emissions or long-term transport. Source tracking analysis by the backward trajectory indicated that the air mass of polluted fog episodes came largely from northern and western China, moved east through Shanxi, Henan, Hebei province to the study area, or from outer Mongolia, crossing the

Jingjinji area to Mt.Tai, The passed areas were notable as heavy industry region with frequent coal mining activities and serious pollution. Moreover, the large population and agricultural activities resulted in the numerous pathogenic microorganisms from human or animal fecal dispersed in the atmosphere. Moreover, a small part from southern China, and the passed region were rich of water resources, e.g., Dongting Lake, Huaihe river, Yangtze river etc. The aquatic bacteria such as *Haliscomenobacter* dispersed in the atmosphere typically derived from the evaporation of lakes and rivers water (Figure 7). In contrary, air mass of non-polluted fog episodes originated mostly from the southern China, the marine sources bacteria (Flavobacteria) indicated the release of prokaryotes into the air from sea-air interactions and airborne marine bacteria can be transported to inland through long-term transport.

In addition, wind rose diagram suggest the predominant north wind during polluted fog episodes and wind speeds ranged 1.2-1.6 m/s, whereas in non-polluted fog episodes it was mainly south wind with higher wind speed (2.1-3.5 m/s) than the polluted fog episodes. Wind direction and speed are important meteorological factors influencing fog formation (Fu et al., 2014). Recent studies also indicate the variation of bacterial concentration and community structure conducted by wind (Evans et al., 2006; Jones & Harrison, 2004). In the present study, air mass from the contaminated area through long-term transport or local regional emission combined with lower wind speeds, largely reduce the diffusion rate of pollutants and thus lead to the sustained high $PM_{2.5}$ during polluted fog episodes. Whereas in the non-polluted fog episodes, higher wind speed was beneficial to the diffusion of pollutants and resulted in the lower $PM_{2.5}$ mass concentration, which finally led to the significant discrepancy of bacterial community structure. However, further research still needed to address the detailed interaction between bacterial community and environmental factors, and understanding the mechanism that how chemical composition influence microbial community.

4. Conclusion

In summary, this work on fog water provided a thorough investigation on bacterial ecological diversity under polluted and non-polluted fog episodes, enhanced understanding the distribution and dispersion of bacteria and their potential

involvements in ecosystem variation and human health. To some degree, $PM_{2.5}$ seams a pivotal variable in shaping bacterial community, which is likely to provide a more comprehensive understanding of the factors controlling the atmospheric water biodiversity under environmental stress. These results provide a basic understanding of mechanism of bacterial community response and metabolism in polluted weather for further study.

Acknowledgement

This work was supported by National Natural Science Foundation of China (41605113, 41375126), Taishan Scholar Grand (ts20120552), China Postdoctoral Science Foundation (No. 2015M582095).

Reference

Abdelelhaleem, D.: *Acinetobacter*: environmental and biotechnological applications, Afr. J. Biotechnol., 4, 71-74, 2003.

Adhikari, A., Reponen, T., Grinshpun, S.A., Dainius. Martuzevicius, LeMasters, G.: Correlation of ambient inhalable bioaerosols with particulate matter and ozone: A two-year study, Environ. Pollut., 140, 16-28, 2006.

Aikawa, M., Hiraki, T., Shoga, M., Tamaki, M.: Fog and precipitation chemistry at Mt. Rokko in Kobe, April 1997-March 1998, Water. Air.Soil. Poll., 130, 1517-1517, 2001.

Alavi, P., Starcher, M.R., Thallinger, G.G., Zachow, C., Müller, H., Berg, G.: *Stenotrophomonas* comparative genomics reveals genes and functions that differentiate beneficial and pathogenic bacteria, BMC Genomics, 15, 1-15, 2013.

Amato, P., Demeer, F., Melaouhi, A., Fontanella, S., Martin-Biesse, A.-S., Sancelme, M., et al.: A fate for organic acids, formaldehyde and methanol in cloud water:their biotransformation by micro-organisms, Atmos. Chem. Phys., 7, 4159-4169, 2007a.

Amato, P., Joly, M., Schaupp, C., Attard, E., Möhler, O., Morris, C.E., et al.: Survival and ice nucleation activity of bacteria as aerosols in a cloud simulation chamber, Atmos. Chem. Phys., 15, 6455-6465, 2015.

Amato, P., Ménager, M., Sancelme, M., Laj, P., Mailhot, G., Delort, A.-M.: Microbial population in cloud water at the Puy de Dôme: Implications for the chemistry of clouds, Atmos. Environ., 39, 4143-4153, 2005.

Amato, P., Parazols, M., Sancelme, M., Laj, P., Mailhot, G., Delort, A.M.: Microorganisms isolated from the water phase of tropospheric clouds at the Puy de Dome: major groups and growth abilities at low temperatures, FEMS. Microbiol. Ecol., 59, 242-254, 2007b.

Amato, P., Parazols, M., Sancelme, M., Mailhot, G., Laj, P., Delort, A.-M.: An important oceanic source of micro-organisms for cloud water at the Puy de Dôme (France), Atmos. Environ., 41, 8253-8263, 2007c.

Anderson, M.J., Willis, T.J.: Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology, Ecology., 84, 511-525, 2008.

Badarinath, K.V.S., Latha, K.M., Chand, T.R.K., Reddy, R.R., Gopal, K.R., Reddy, L.S.S., et al.: Black carbon aerosols and gaseous pollutants in an urban area in North India during a fog period, Atmos. Res., 85, 209-216, 2007.

Banks, L.N., Kurdy, N.M., Hassan, I., Aster, A.S.: Septic arthritis of the ankle due to *Moraxella osloensis*, Foot. Ankle. Surg., 13, 189-191, 2007.

Bock, C., Kroppenstedt, R.M., Diekmann, H.: Degradation and bioconversion of aliphatic and aromatic hydrocarbons by Rhodococcus ruber 219, Appl. Microbiol. Biot., 45, 408-410, 1996.

Boonchan, S., Britz, M.L., Stanley, G.A.: Surfactant-enhanced biodegradation of high molecular weight polycyclic aromatic hydrocarbons by *Stenotrophomonas maltophilia*, Biotechnol. Bioeng., 59, 482-494, 1998.

Boris, A.J., Lee, T., Park, T., Choi, J., Seo, S., Collett, J.L.: Fog composition at Baengnyeong Island in the Eastern Yellow Sea: detecting markers of aqueous atmospheric oxidations, Atmos. Chem. Phys., 15, 24871-24908, 2015.

Bottone, E.J., Madayag, R.M., Qureshi, M.N.: Acanthamoeba keratitis: synergy between amebic and bacterial cocontaminants in contact lens care systems as a prelude to infection, J. Clin. Microbiol., 30, 2447-2450, 1992.

Bowers, R.M., Clements, N., Emerson, J.B., Wiedinmyer, C., Hannigan, M.P., Fierer, N.: Seasonal variability in bacterial and fungal diversity of the near-surface atmosphere, Environ. Sci. Technol., 47, 12097-12106, 2013.

Bowers, R.M., Lauber, C.L., Wiedinmyer, C., Hamady, M., Hallar, A.G., Fall, R., et al.: Characterization of airborne microbial communities at a high-elevation site and their potential to act as atmospheric ice nuclei, Appl. Environ. Microbiol., 75, 5121-5130, 2009.

Burrows, S.M., Elbert, W., Lawrence, M.G., Pöschl, U.: Bacteria in the global atmosphere-part 1: review and synthesis of literature data for different ecosystems, Atmos. Chem. Phys., 9, 9263-9280, 2009.

Busse, H., Denner, E., Buczolits, S., Salkinoja-Salonen, M., Bennasar, A., Kämpfer, P.: *Sphingomonas aurantiaca* sp. nov., *Sphingomonas aerolata* sp. nov. and *Sphingomonas faeni* sp. nov., air- and dustborne and Antarctic, orange-pigmented, psychrotolerant bacteria, and emended description of the genus *Sphingomonas*, Int. J. Syst. Evol. Micr., 53, 1253-1260, 2003.

Cao, C., Jiang, W., Wang, B., Fang, J., Lang, J., Tian, G., et al.: Inhalable microorganisms in Beijing's $PM_{2.5}$ and PM_{10} pollutants during a severe smog event, Environ. Sci. Technol., 48, 1499-1507, 2014.

Cho, B.C., Jang, G.I.: Active and diverse rainwater bacteria collected at an inland site in spring and summer 2011, Atmos. Environ., 94, 409-416, 2014.

Clarke, K.R.: Non-parametric multivariate analyses of changes in community structure, Austral Ecol, 18, 117-143, 1993.

Delort, A.M., Vaïtilingom, M., Amato, P., Sancelme, M., Parazols, M., Mailhot, G., et al.: A short overview of the microbial population in clouds: Potential roles in atmospheric chemistry and nucleation processes, Atmos. Res., 98, 249-260, 2010.

Després, V.R., Alex Huffman, J., Burrows, S.M., Hoose, C., Safatov, A.S., Buryak, G., et al.: Primary biological aerosol particles in the atmosphere: a review, Tellus. B., 64, 145-153, 2012.

Dong, L., Qi, J., Shao, C., Zhong, X., Gao, D., Cao, W., et al.: Concentration and size

distribution of total airborne microbes in hazy and foggy weather, Sci. Total. Environ., 541, 1011-1018, 2016.

Donn, S., Kirkegaard, J.A., Perera, G., Richardson, A.E., Watt, M.: Evolution of bacterial communities in the wheat crop rhizosphere, Environ. Microbiol., 17, 610–621, 2015.

Evans, C.A., Coombes, P.J., Dunstan, R.H.: Wind, rain and bacteria: The effect of weather on the microbial composition of roof-harvested rainwater, Water. Res., 40, 37-44, 2006.

Fernández-González, R., Yebra-Pimentel, I., Martínez-Carballo, E., Simal-Gándara, J., Pontevedra-Pombal, X.: Atmospheric pollutants in fog and rain events at the northwestern mountains of the Iberian Peninsula, Sci.Total. Environ., 497–498, 188-199, 2014.

Fu, G.Q., Xu, W.Y., Yang, R.F., Li, J.B., Zhao, C.S.: The distribution and trends of fog and haze in the North China Plain over the past 30 years, Atmos. Chem. Phys., 14, 11949-11958, 2014.

Fuzzi, S., Mandrioli, P., Perfetto, A.: Fog droplets-an atmospheric source of secondary biological aerosol particles, Atmos. Environ., 31, 287-290, 1997.

Gandolfi, I., Bertolini, V., Bestetti, G., Ambrosini, R., Innocente, E., Rampazzo, G., et al.: Spatio-temporal variability of airborne bacterial communities and their correlation with particulate matter chemical composition across two urban areas, Appl. Microbiol. Biot., 99, 4867-4877, 2015.

Gao, M., Qiu, T., Jia, R., Han, M., Song, Y., Wang, X.: Concentration and size distribution of viable bioaerosols during non-haze and haze days in Beijing, Environmental Science and Pollution Research, 22, 4359-4368, 2015.

Gao, M., Yan, X., Qiu, T., Han, M., Wang, X.: Variation of correlations between factors and culturable airborne bacteria and fungi, Atmos. Environ., 128, 10-19, 2016.

Geng, L., Ming, C., Liang, Q., Wei, L., Wei, Z., Ping, S., et al.: Functional analysis of a putative regulatory gene, tadR , involved in aniline degradation in Delftia tsuruhatensis AD9, Arch. Microbiol., 191, 603-614, 2009.

Gilad, J., Borer, A., Peled, N., Riesenberg, K., Tager, S., Appelbaum, A., et al.: Hospital-acquired brevundimonas vesicularis septicaemia following open-heart surgery: case report and literature review, Infect. Dis., 32, 90-91, 2009.

Gonzalezbashan, L.E., Lebsky, V.K., Hernandez, J.P., Bustillos, J.J., Bashan, Y.: Changes in the metabolism of the microalga *Chlorella vulgaris* when coimmobilized in alginate with the nitrogen-fixing *Phyllobacterium myrsinacearum*, Can. J. Microbiol., 46, 653-659, 2000.

Goyal, A.K., Zylstra, G.J.: Molecular cloning of novel genes for polycyclic aromatic hydrocarbon degradation from Comamonas testosteroni GZ39, Appl. Environ. Microbiol., 62, 230-236, 1996.

Griffiths, E., Gupta, R.S.: Identification of signature proteins that are distinctive of the Deinococcus-Thermus phylum, Int. Microbiol., 10, 201-208, 2007.

Gu, H., Lou, J., Wang, H., Yang, Y., Wu, L., Wu, J., et al.: Biodegradation, Biosorption of Phenanthrene and Its Trans-Membrane Transport by *Massilia* sp. WF1 and *Phanerochaete chrysosporium*, Front. Microbiol., 7, e40842, 2016.

Guo, J., Wang, Y., Shen, X., Wang, Z., Lee, T., Wang, X., et al.: Characterization of cloud water chemistry at Mount Tai, China: Seasonal variation, anthropogenic

impact, and cloud processing, Atmos. Environ., 60, 467-476, 2012.
Han, X.Y., Andrade, R.A.: Brevundimonas diminuta infections and its resistance to fluoroquinolones, J. Antimicrob. Chemoth., 55, 853-859, 2005.
Hauser, E., Kämpfer, P., Busse, H.J.: *Pseudomonas psychrotolerans* sp. nov, Int. J. Syst. Evol. Micr., 54, 1633-1637, 2004.
Hill, K.A., Shepson, P.B., Galbavy, E.S., Anastasio, C., Kourtev, P.S., Konopka, A., et al.: Processing of atmospheric nitrogen by clouds above a forest environment, J. Geophys. Res., 112, 2864-2864, 2007.
Huang, K., Zhuang, G., Wang, Q., Fu, J.S., Lin, Y., Liu, T., et al.: Extreme haze pollution in Beijing during January 2013: chemical characteristics, formation mechanism and role of fog processing, Atmos. Chem. Phys. Discuss., 2014, 7517-7556, 2014.
Hugo, C.J., Bruun, B., Jooste, P.J.: The Genera Empedobacter and Myroides, in, ed., edited by, Springer New York, 630-637, 2005.
Jones, A.M., Harrison, R.M.: The effects of meteorological factors on atmospheric bioaerosol concentrations-a review, Sci. Total. Environ., 326, 151-180, 2004.
Kämpfer, P., Lodders, N., Huber, B., Falsen, E., Busse, H.J.: *Deinococcus aquatilis* sp. nov., isolated from water, Int. J. Syst. Evol. Micr., 58, 2803-2806, 2009.
Katra, I., Arotsker, L., Krasnov, H., Zaritsky, A., Kushmaro, A., Ben-Dov, E.: Richness and diversity in dust stormborne biomes at the southeast mediterranean, Scientific reports, 4, 5265, 10.1038/srep05265, 2014.
Kopcakova, A., Stramova, Z., Kvasnova, S., Godany, A., Perhacova, Z., Pristas, P.: Need for database extension for reliable identification of bacteria from extreme environments using MALDI TOF mass spectrometry, Chem. Pap., 68, 1435-1442, 2014.
Kourtev, P.S., Hill, K.A., Shepson, P.B., Konopka, A.: Atmospheric cloud water contains a diverse bacterial community, Atmos. Environ., 45, 5399-5405, 2011.
Kuczynski, J., Stombaugh, J., Walters, W.A., González, A., Caporaso, J.G., Knight, R.: Using QIIME to analyze 16S rRNA gene sequences from microbial communities:Current Protocols in Microbiology, Curr. Protoc. Bioinformatics., 10, 1-28, 2011.
Masoud, W., Takamiya, M., Vogensen, F.K., Lillevang, S., Al-Soud, W.A., Sørensen, S.J., et al.: Characterization of bacterial populations in Danish raw milk cheeses made with different starter cultures by denaturating gradient gel electrophoresis and pyrosequencing, Int. Dairy. J., 21, 142-148, 2011.
Mattimore, V., Battista, J.R.: Radioresistance of *Deinococcus radiodurans*: functions necessary to survive ionizing radiation are also necessary to survive prolonged desiccation, J. Bacteriol., 178, 633-637, 1996.
McIlroy, S.J., Nielsen, P.H.: The Family Saprospiraceae, in: The Prokaryotes-Other Major Lineages of Bacteria and the Archaea, ed., edited by, E. Rosenberg et al. (eds.), 863-889, 2014.
Mohan, M., Payra, S.: Influence of aerosol spectrum and air pollutants on fog formation in urban environment of megacity Delhi, India, Environ. Monit. Assess., 151, 265-277, 2009.
Mortazavi, R., Attiya, S., Ariya, P.A.: Arctic microbial and next-generation sequencing approach for bacteria in snow and frost flowers: selected identification, abundance and freezing nucleation, Atmos. Chem. Phys., 15, 6183-6204, 2015.

Nemec, A., De Baere, T., Tjernberg, I., Vaneechoutte, M., van der Reijden, T.J., Dijkshoorn, L.: *Acinetobacter ursingii* sp. nov. and *Acinetobacter schindleri* sp. nov., isolated from human clinical specimens, Int. J. Syst. Evol. Micr., 51, 1891-1899, 2001.

Newton, R.J., Jones, S.E., Alexander, E., Mcmahon, K.D., Stefan, B.: A guide to the natural history of freshwater lake bacteria, Microbiol. Molecul. Biol. R., 75, 14-49, 2011.

Nováková, D., Sedlácek, I., Pantůcek, R., Stetina, V., Svec, P., Petrás, P.: *Staphylococcus equorum* and *Staphylococcus succinus* isolated from human clinical specimens, J. Med. Microbiol., 55, 523-528, 2006.

Pérez-Pantoja, D., Donoso, R., Agulló, L., Córdova, M., Seeger, M., Pieper, D.H., et al.: Genomic analysis of the potential for aromatic compounds biodegradation in Burkholderiales, Environ. Microbiol., 14, 1091–1117, 2012.

Poeschl, U.: Atmospheric aerosols: composition, transformation, climate and health effects, Angewandte Chemie International Edition, 37, 7520-7540, 2006.

Saddler, G.S., Bradbury, J.F.: Xanthomonadales ord. nov, in: Bergey's Manual® of Systematic Bacteriology, ed., edited by, Springer US, 63-122, 2005.

Schloss, P., Westcott, S., Ryabin, T., Hall, J., Hartmann, M., Hollister, E., et al.: Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities, Appl. Environ. Microbiol., 75, 7537-7541, 2009.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., et al.: Metagenomic biomarker discovery and explanation, Genome. Biol., 12, 1-18, 2011.

Seo, Y.S., Lee, D.Y., Rayamahji, N., Mi, L.K., Han, S.Y.: Biofilm-forming associated genotypic and phenotypic characteristics of Staphylococcus spp. isolated from animals and air, Res. Vet. Sci., 85, 433-438, 2008.

Stanier, R.Y., Palleroni, N.J., Doudoroff, M.: The aerobic pseudomonads: a taxonomic study, J. Gen. Microbiol., 43, 159-271, 2010.

Suominen, I., Andersson, M.A., Andersson, M.C., Hallaksela, A.-M., Kämpfer, P., Rainey, F.A., et al.: Toxic *Bacillus pumilus* from indoor air, recycled paper pulp, Norway spruce, food poisoning outbreaks and clinical samples, Syst. Appl. Microbiol., 24, 267-276, 2001.

Trilla, A., Miro, J.M.: Identifying high risk patients for *Staphylococcus aureus* infections: skin and soft tissue infections, J. Chemoth, 7 Suppl 3, 37-43, 1995.

Vaïtilingom, M., Amato, P., Sancelme, M., Laj, P., Leriche, M., Delort, A.M.: Contribution of microbial activity to carbon chemistry in clouds, Appl. Environ. Microbiol., 76, 23-29, 2010.

Vaïtilingom, M., Attard, E., Gaiani, N., Sancelme, M., Deguillaume, L., Flossmann, A.I., et al.: Long-term features of cloud microbiology at the puy de Dôme (France), Atmos. Environ., 56, 88-100, 2012.

Vaïtilingom, M., Deguillaume, L., Vinatier, V., Sancelme, M., Amato, P., Chaumerliac, N., et al.: Potential impact of microbial activity on the oxidant capacity and organic carbon budget in clouds, Proc. Natl. Acad. Sci. USA., 110, 559-564, 2013.

Xu, P., Yu, B., Li, F.L., Cai, X.F., Ma, C.Q.: Microbial degradation of sulfur, nitrogen and oxygen heterocycles, Trends. Microbiol, 14, 398-405, 2006.

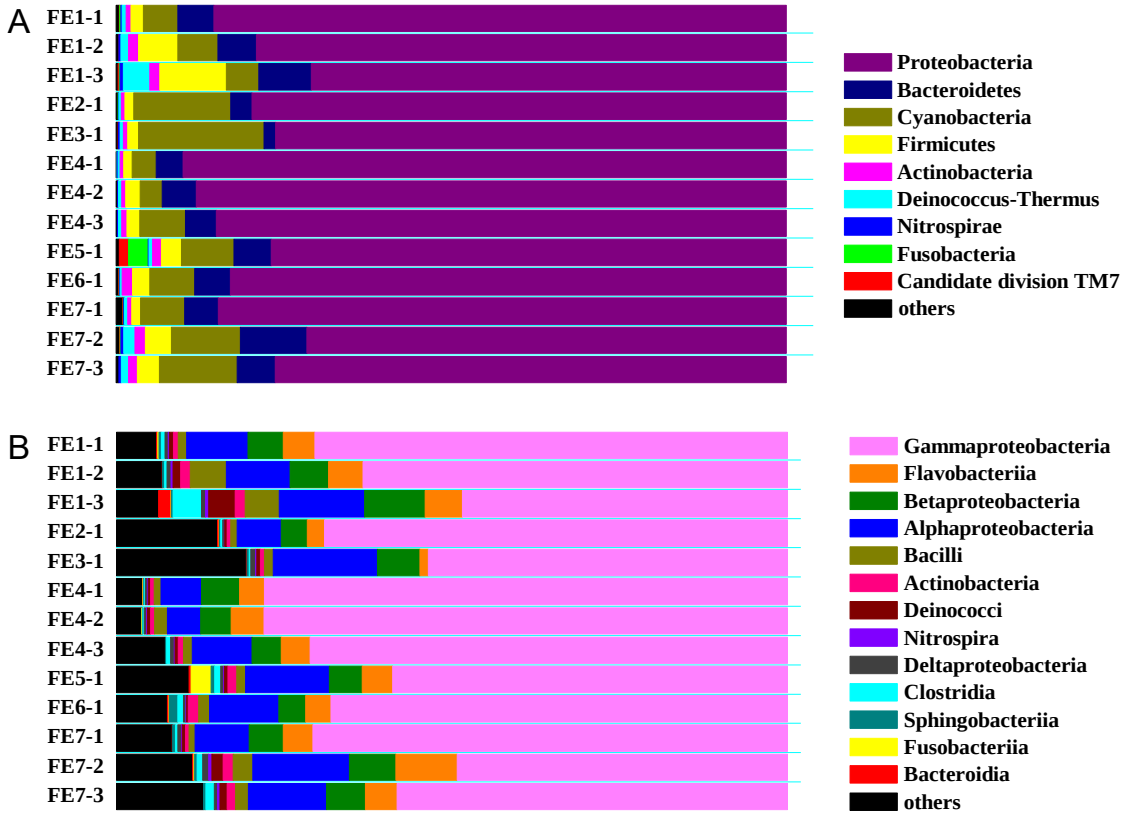


Figure 1 Bacterial community variation for the fog episodes at the phylum and class level. Predominant taxa higher than 1% are indicated in the bar graphs.

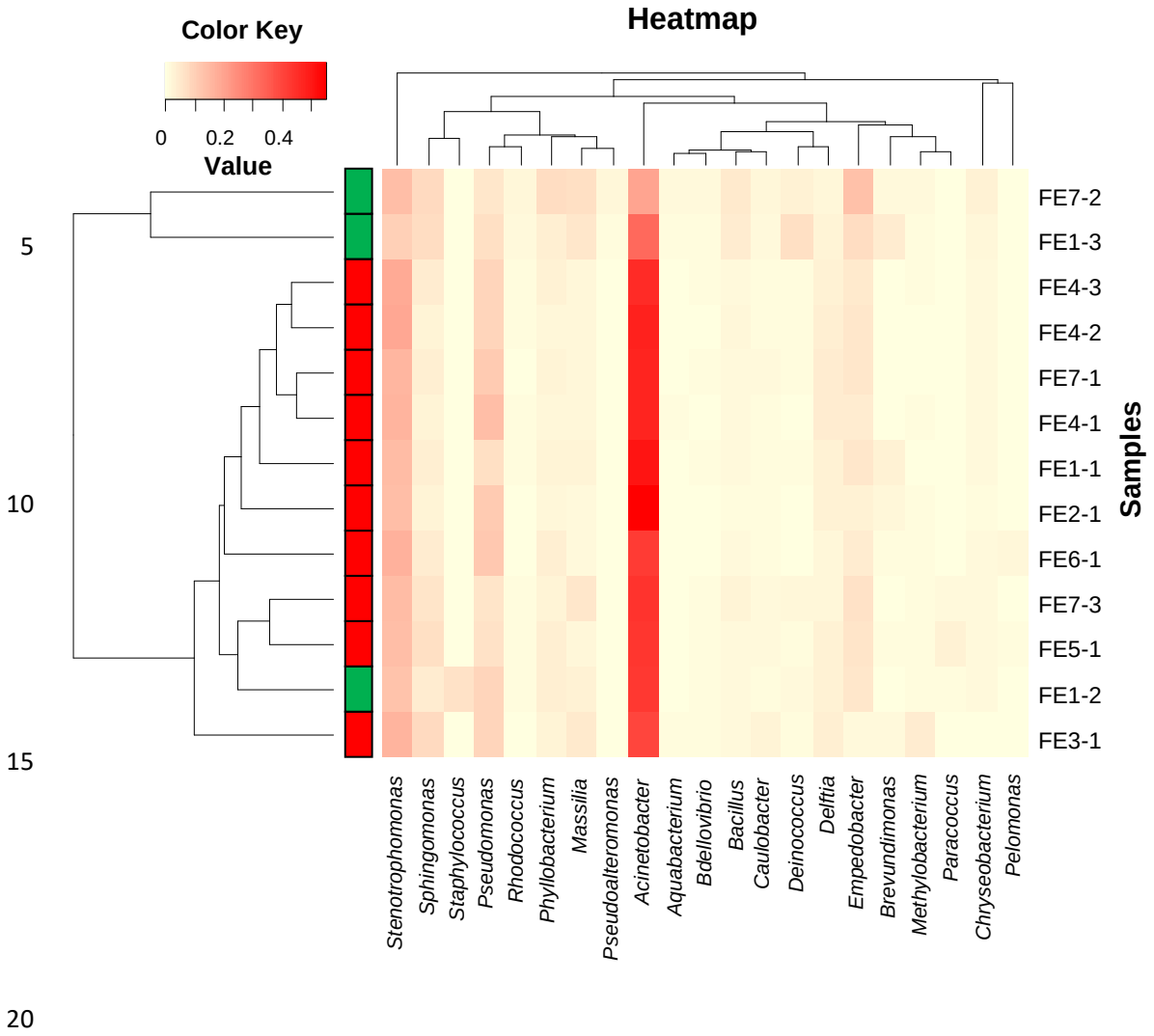


Figure 2 Hierarchically clustered heatmap of the predominant bacterial genus distribution under polluted and non-polluted fog episodes. Polluted fog water samples are indicated by red square, non-polluted samples are green.

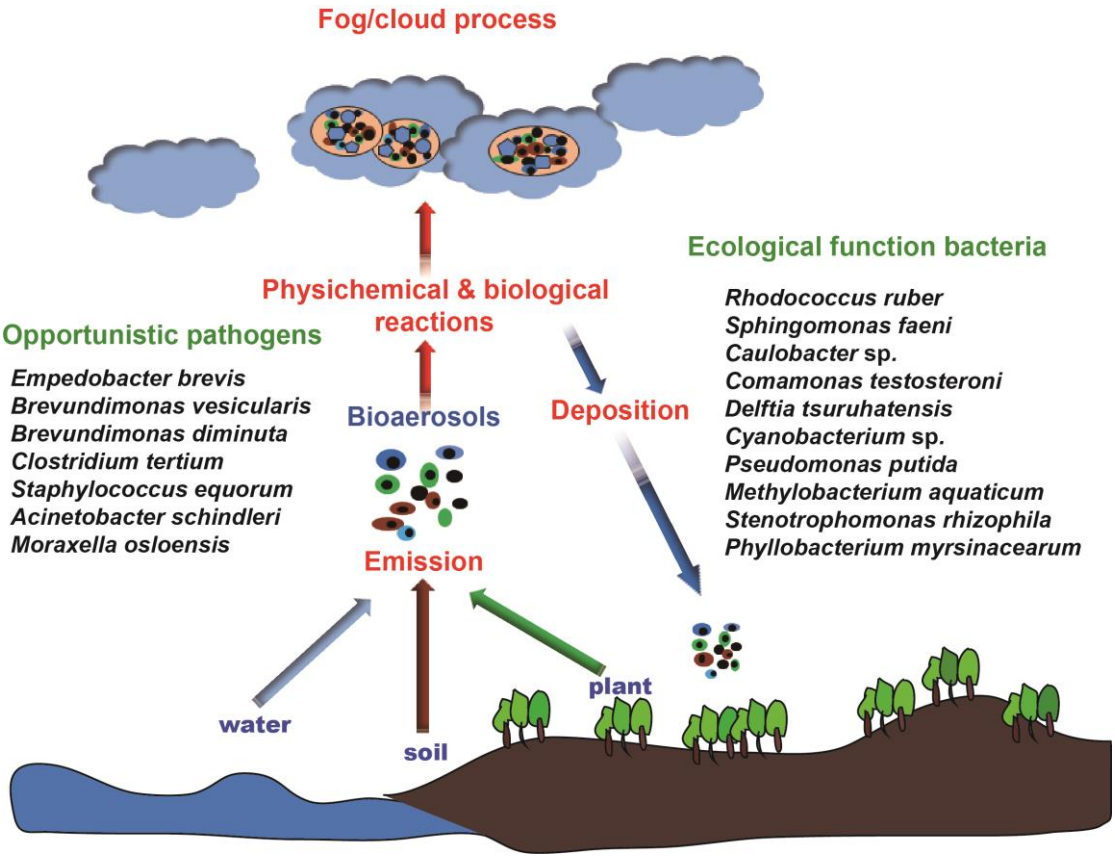


Figure 3 Schematic representation of the life cycle and potential influence on the ecosystem of bioaerosols in the atmosphere, modified from Poeschl (Poeschl, 2006). The predominant identified bacteria species with potential ecological functions are indicated in the figure.

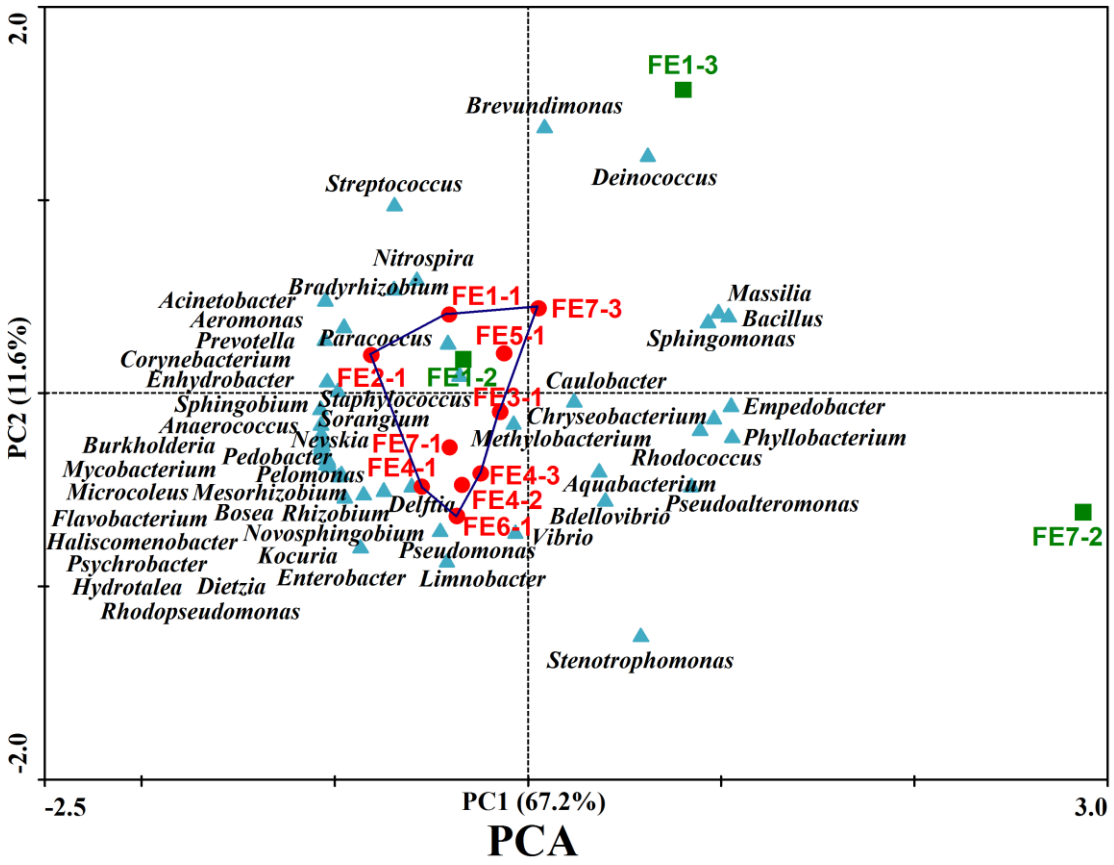


Figure 4 Principal component analysis shows the bacterial community variability between polluted and non-polluted fog episodes. Samples in the same group indicate the cluster similarity.

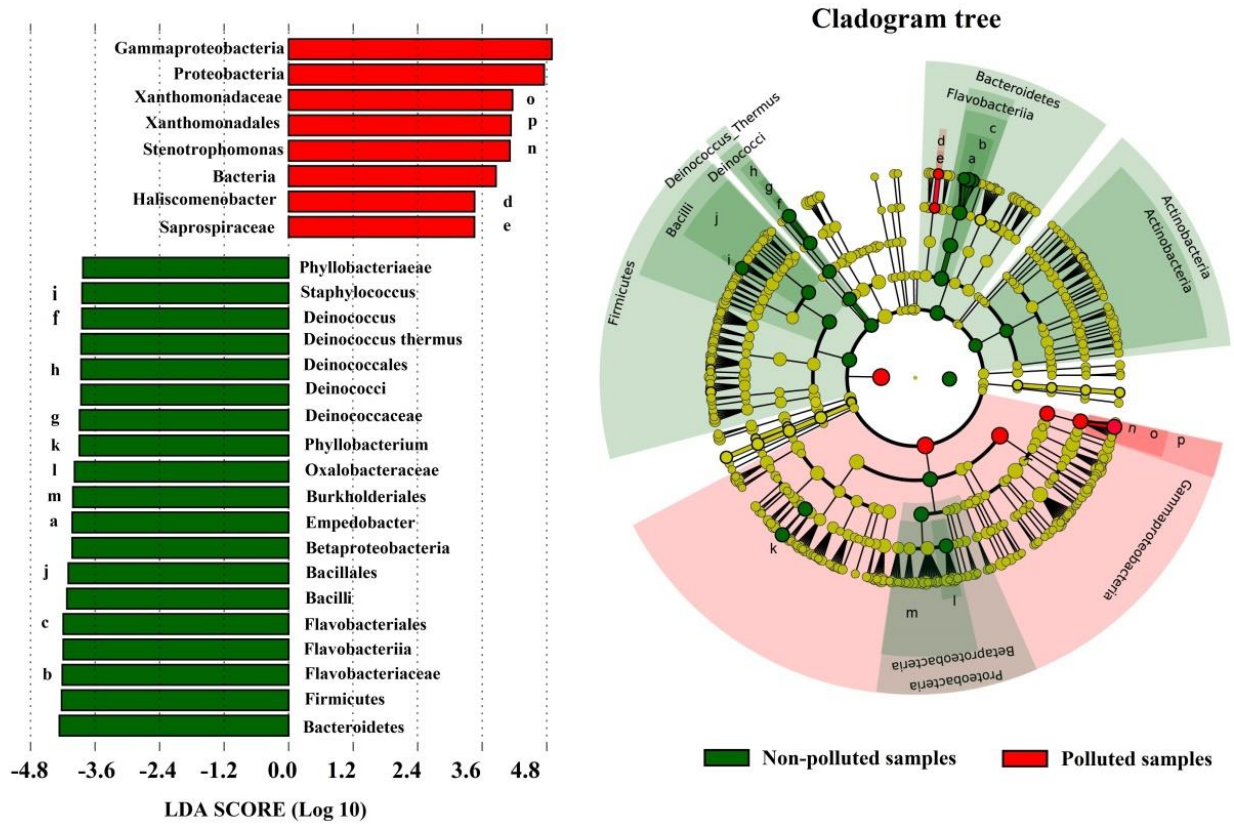


Figure 5 Cladogram of phylogenetic relationships of bacterial lineages associated with polluted and non-polluted fog episodes; taxa with LDA values higher than 3.5 by LEfSe are displayed. Differences are represented with different color (red indicating polluted fog episodes, green non-polluted fog episodes, and yellow nonsignificant).

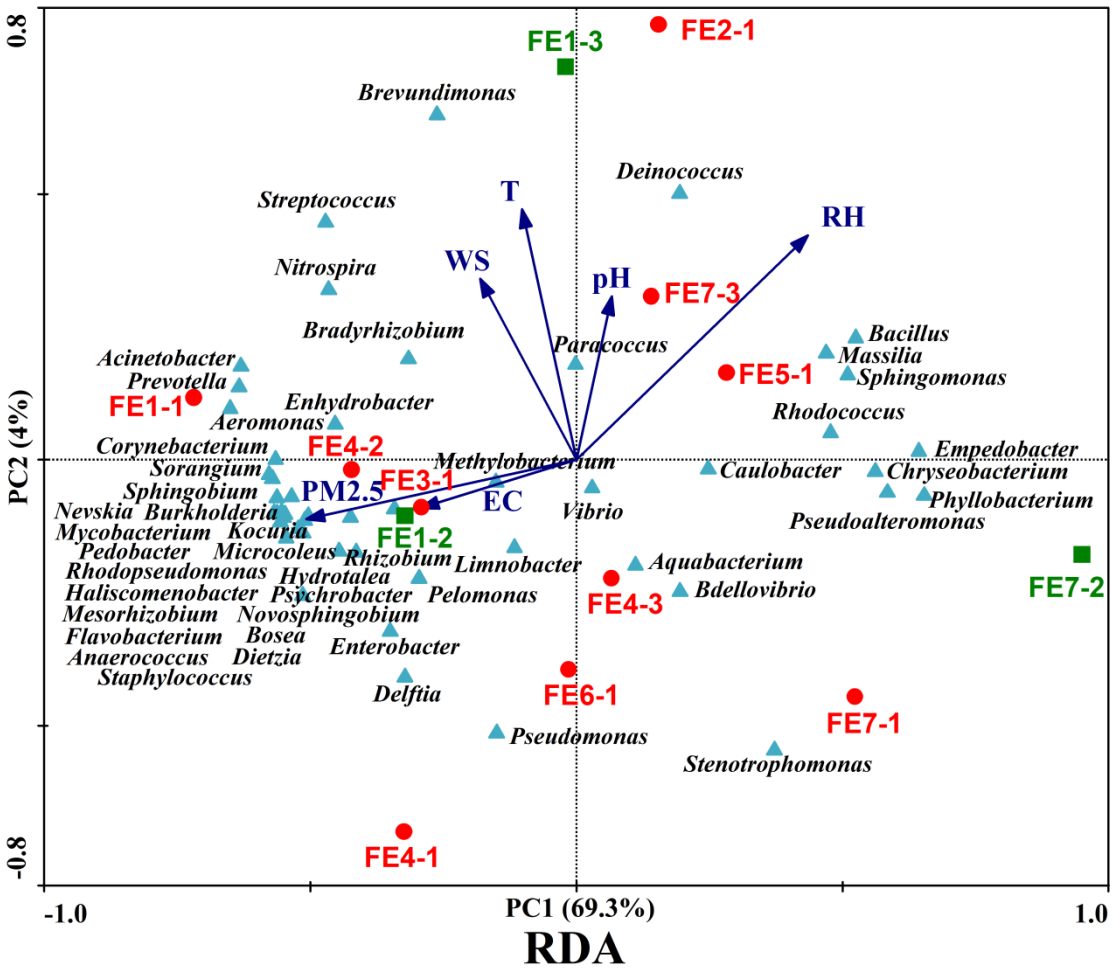


Figure 6 Biplot of the environmental variables and predominant genera using a redundancy analysis (RDA) model. Species data are listed in Table S2. The selected environmental variables are significant ($P < 0.05$) using Monte Carlo permutation testing. Species are labeled with triangle, the closer of which indicates existence in similar environment. Environmental variables are showed by arrows; the relative length is positive correlation with the importance in influencing bacterial community structure. The angle between the arrow and the ordination axis suggest the variable response respect to the RDA gradient.

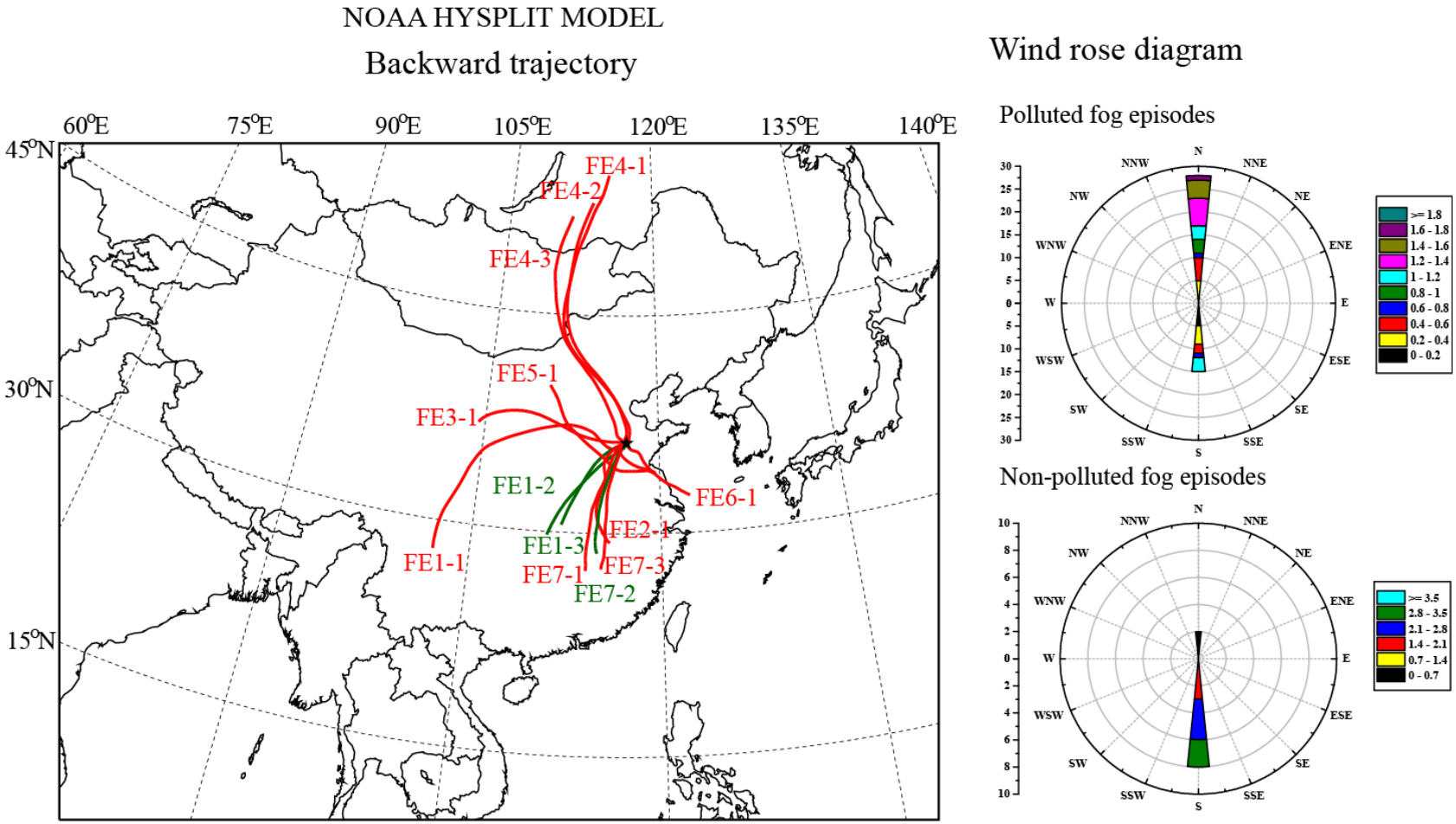


Figure 7 Air mass transport pathways for the fog episodes using the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model. 24-hour backward trajectories were calculated for air parcels arriving at 1534 m above sea level. Wind Rose Diagram of Study Area during sampling time.

Table 1 Description fog episodes at Mt.Tai, China

Fog episodes	Data	Samples	Start time (BJT)	Stop time (BJT)	Duration (h)	$PM_{2.5}$ ($\mu g \cdot m^{-3}$)	pH	EC ($\mu S \cdot cm^{-1}$)	Pollution
FE1	24 Jul 2014	FE1-1	8:50	15:30	6:40	105.07	4.03	583	A
		FE1-2	15:30	17:30	2:00	22.35	4.32	219.2	B
		FE1-3	17:30	22:51	5:21	14.66	5.74	104.4	B
FE2	5 Aug 2014	FE2-1	6:45	9:17	2:32	30.36	5.80	275.7	A
FE3	5 Aug - 6 Aug 2014	FE3-1	19:05	4:01	8:56	42.25	5.10	501	A
FE4	14 Aug - 15 Aug 2014	FE4-1	22:41	0:44	2:03	42.69	6.36	170.4	A
		FE4-2	0:44	5:06	4:22	47.98	5.34	86.34	A
		FE4-3	5:06	6:03	0:57	36.88	4.89	64.95	A
FE5	17 Aug 2014	FE5-1	10:10	11:18	1:08	63.18	5.20	120.5	A
FE6	17 Aug - 18 Aug 2014	FE6-1	22:18	1:25	3:07	54.33	3.80	321.8	A
FE7	23 Aug 2014	FE7-1	2:30	4:38	2:08	30.45	4.38	356.2	A
		FE7-2	4:38	6:21	1:43	23.39	5.01	207.5	B
		FE7-3	6:21	9:20	2:59	41.60	5.74	187.6	A

BJT refers to Beijing Time, which equals UTC + 8. EC refers to the electric conductivity.
The A, B refers to the the polluted and non-polluted samples based on the WHO 24-hr average standard $PM_{2.5}$ mass concentration ($PM_{2.5}$ = 25 $\mu g \cdot m^{-3}$), respectively.

Table 2 Summary of bacterial diversity and richness of fog water

Sample ID	Reads	OTUs	Ace	Chao1	Coverage	Shannon	Simpson
FE1-1	18213	975	1835	1491	0.9761	3.9418	0.0646
FE1-2	18702	1184	1841	1730	0.9719	4.1919	0.0620
FE1-3	17662	1173	1689	1687	0.9732	4.7067	0.0327
FE2-1	19914	1125	1756	1684	0.9744	3.9582	0.0630
FE3-2	18199	1022	2082	1582	0.9734	3.9749	0.0647
FE4-1	18350	941	1828	1461	0.9762	3.6041	0.0953
FE4-2	17707	967	1522	1427	0.9752	3.6748	0.0902
FE4-3	17397	981	2091	1611	0.9725	3.8074	0.0832
FE5-1	16384	1132	1814	1790	0.9676	4.3173	0.0546
FE6-1	16896	1186	1997	1872	0.9657	4.1268	0.0666
FE7-1	16350	1103	2501	1795	0.965	3.9040	0.0810
FE7-2	18252	1150	1732	1673	0.9729	4.3709	0.0426
FE7-3	18122	1258	1958	1999	0.9686	4.3776	0.0531
Aerosol (Katra et al., 2014)	4020	1412		2142	0.8300		
Rain water in July (Cho & Jang, 2014)	3055	1542	13083	6387			
$PM_{2.5}$ in summer (Franzetti et al., 2011)		2222		4036			

The diversity indexes including OTUs, ace, Chao1, coverage, shannon, simpson were defined at 97% sequence similarity.

Table 3 The predominate bacterial species identified in fog water samples

Species	Habitat	Ecological roles	Polluted (%)	Non-polluted (%)
Empedobacter brevis[BA]	soil/water/plant	Opportunistic pathogens	2.8796	4.8345
Brevundimonas vesicularis[BP]	soil/water	Opportunistic pathogens	0.2126	0.8271
Brevundimonas diminuta[BP]	soil/water	Opportunistic pathogens	0.2811	0.0579
Clostridium tertium[FR]	soil/faeces	Opportunistic pathogens	0.0269	0.0055
Staphylococcus equorum[GP]	soil/water/annimal	Opportunistic pathogens	0.0044	0.0089
Acinetobacter schindleri[GP]	soil/water	Opportunistic pathogens	0.0084	0.0238
Moraxella osloensis[GP]	soil/animal	Opportunistic pathogens	0.097	0.0962
Rhodococcus ruber[AC]	soil/water	Biodegradation	0.3314	0.6226
Chryseobacterium aquaticum[BA]	soil/water	Protect and promote plants growth	0.0307	0.073
Chryseobacterium jejuense[BA]	soil/water	Protect and promote plants growth	0.6698	1.0107
Deinococcus aquatilis[DT]	soil/water	Extremophiles, radiation-resistant	0.0193	0.036
Novosphingobium aromaticivorans[AP]	soil/water	Biodegradation	0.2178	0.2007
Sphingomonas faeni[AP]	soil/water	Biodegradation	0.0659	0.1799
Sphingomonas kaistensis[AP]	soil/water	Biodegradation	0.0529	0.0417
Sphingomonas leidyi[AP]	soil/water	Biodegradation	0.0305	0.0149
Methylobacterium aquaticum[AP]	water	Methylotrophic, carbon fixing	0.0915	0.0955
Methylobacterium adhaesivum[AP]	soil/water	Methylotrophic, carbon fixing	0.0574	0.0808

Caulobacter sp.[AP]	water	Biodegradation	0.6916	0.7264
Phyllobacterium myrsinacearum[AP]	soil/plant	Rhizosphere bacteria, nitrogen fixation	1.3408	2.3989
Comamonas testosteroni[BP]	soil/water	Biodegradation	0.0788	0.0502
Delftia tsuruhatensis[BP]	soil/water	Biodegradation	1.9164	1.3085
Cyanobacterium sp.[CY]	soil/water	Carbon and nitrogen fixing	1.8229	0.1444
Pseudomonas geniculata[GP]	soil/water/plant	Biodegradation	0.0516	0.0293
Pseudomonas putida[GP]	water/soil	Biodegradation/protect and promote plants growth	0.0263	0.0457
Pseudomonas psychrotolerans[GP]	soil/water	Extremophiles, psychrotolerant	0.0324	0.0217
Stenotrophomonas rhizophila[GP]	soil/water/plant	Rhizosphere bacteria, plant-protective	0.6016	1.1263

Abbreviates are as followed: AP, Alphaproteobacteria; BP, Betaproteobacteria; GP, Gammaproteobacteria; AC, Actinobacteria; BA, Bacteroidetes; CY, Cyanobacteria; DT, Deinococcus-Thermus; FR, Firmicutes. Biodegradation refers to the bacteria associated with the biodegradation of organic compounds, even toxic pollutants, e.g., aromatic compounds.