Artificial Intelligence, Medical Engineering and Education
Z.B. Hu et al. (Eds.)

doi:10.3233/ATDE231352

Investigation into the Long-Term Dynamics of Microbiomes in Hong Kong Outdoor Bioaerosols

Lingxiao XU[a,1], Xinzhao TONG[b]
[a] *Faculty of Natural Sciences, Imperial College London, London, UK*
b Department of Biological Sciences, Xi'an Jiaotong-Liverpool University, Suzhou, China

Abstract. The composition of outdoor microorganisms is dynamic in nature, often affected by factors such as geography, seasonality, environmental conditions. In this study, bioinformatic platform QIIME2 (v2022.8) was used to process the 93 outdoor air samples after 16S rRNA gene sequencing, aiming to understand the dynamics of the outdoor bioaerosol microbial communities in Hong Kong and identify the significant factor underlying such variations. The results suggested the diversity and composition of outdoor microorganisms in Hong Kong significantly varied by seasonal changes, with the outdoor air bacterial community being highest in summer and autumn and lowest in winter. The study also identified indicator species for the three seasons, spring, autumn and winter, to investigate the changes in bacterial habits, living conditions and environment during the different seasons. The results indicate that bacteria in the aquatic environment are more abundant during these seasons, and that spring has seen the emergence of bacteria often associated with human epidemics. However, within one season, geographic area had minimal influence on the bacterial community diversity and composition, suggesting outdoor airborne bacterial communities tended to be homogeneous across different locations on a city-wide scale within a season.

Keywords. Metagenomics; Microbial ecology; Environmental microbiology.

1.Introduction

In several cases in recent years, the experience of major infectious disease events has reinforced the conviction of researchers that exploring the substance of airborne microorganisms has a significant impact on the prevention and interdiction of pathogenic infectious agents. The major chemical constituents of outdoor air include particulate matter 2.5 (PM2.5) and the organisms, such as bacteria, fungi, and viruses [1]. Airborne pathogenic microorganisms may cause human respiratory infection after inhalation, especially during the cold and wet seasons. Hence, understanding the dynamic pattern of the outdoor bioaerosol microbiota lays a foundation for the prevention of the spread of outdoor infections and prediction of large-scale transmission of diseases in outdoor urban places [2].

Seasonality and geography are the two major factors significantly affecting the diversity and composition of outdoor airborne microbial communities. In most regions

[1] Corresponding Author: Lingxiao XU, E-mail: lx323@ic.ac.uk

of the world, the high humidity and temperature in summer favor the growth and reproduction of microorganisms. As a result, summer usually harbors the highest airborne microbial diversity [3]. The influence of geography on the diversity and composition of the atmospheric microbiome varies across different geographical scales, with significant differences detected between continents to even small local scales. However, due to the less vegetation in the city compared to the suburbs and rural areas, urbanization can result in the homogenization of the outdoor airborne microbial communities in urban areas [2].

In this study, the QIIME2(v2022.8) bioinformatics platform [1]was applied to study the diversity and composition of the outdoor airborne bacterial community in Hong Kong across four seasons and using the internal R language packages to calculate and visualize the resulting data, aiming to identify the factors significantly structuring the dynamics pattern of the outdoor airborne microbiome [4].

2. Sample Collection and Sequence Data Analysis

Previous researchers collected air samples for microbiome analysis from Hong Kong outdoor aerosol environment of commercial buildings, including five office buildings and three shopping malls, during the weekends to reflect the normal use of the outdoor environment of high-volume buildings with 3 days in each season and negative control samples. For each sample, three PCR reactions were performed in triplicate. The amplicons were pooled and filtered prior to institution construction, and then sequenced the platform. The 16S rRNA gene fragment sequences required for the experiments were finally obtained that could verify the bacterial species.

The raw sequences were downloaded from NCBI under the accession number PRJNA55682. The samples were grouped according to the seasons and sampling areas when and where they were collected. However, one air sample was removed from the study due to the incompleteness of data upon examination, resulting in a dataset consisting of 22 winter, 23 spring, 24 summer, and 24 autumn samples, plus a negative control sample.

In this study, QIIME2 (v2022.8) was used for bioinformatics analyses [1]. The import command was used to demultiplex and import raw paired-end fastq files into QIlME2. Using the DADA2 denoise-paired plug-in utility, the DADA2 algorithm was executed to generate amplicon sequence variants (ASVs) [2]. with a minimum read length of 250 bp and a maximum of two expected errors per read, the imported paired-end sequences were filtered, truncated, de-noised, and de-replicated. Additionally, chimeric and singleton sequences were eliminated [5].

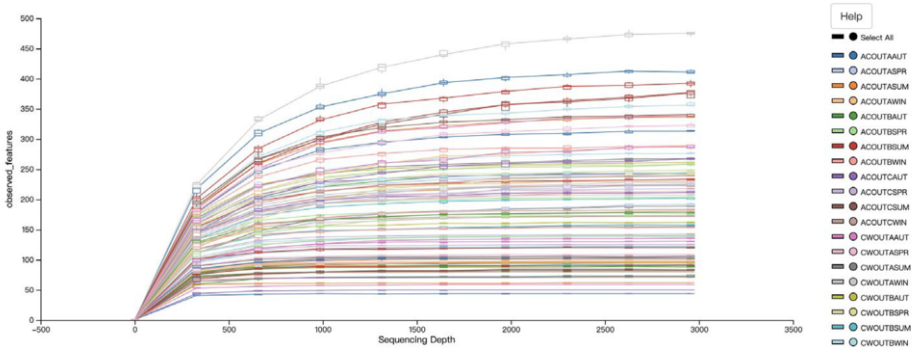

Figure 1. The number of observed sample features at different sequence shows the rarefaction curves.

Using a feature classifier trained with the SILVA database and truncated to the V4 hypervariable region of the 16S rRNA gene, taxonomy was ascribed to all ASVs. Representative sequences were aligned with MAFFT and deployed in QIIME2's FastTree for phylogenetic reconstruction. Mitochondrial, chloroplast, and contaminant sequences removed from all the samples. Based on the blank sample, lineages of taxa with a relative abundance exceeding 5% were deemed contaminants. After contamination filtering, 11,018 unique ASVs were retained for subsequent analysis [2]. After removing the six contaminant samples from the control group, the samples were re-tested for feature analysis again, with the result that when the sequence depth was taken to 2956, the number of discoverable features for essentially all samples would level off and no longer increase, then a very minimal sample size could be kept missing to obtain the lowest level of error value (Figure 1).

At the beginning of the analysis of the data, to obtain the approximate composition of the microbial communities in the four seasons in the most visual and extensive way, R (v3.4.4) packages 'ggplot2' (v3.4.2) and 'dplyr' (v0.7.8) were used to plot histograms with the three taxonomic levels 'phylum', 'class' and 'genus' for the taxonomic classification of the samples in each of the four seasons, providing a visual understanding of the changes in the relative abundance.

The alpha and beta diversity metrics were computed using the core-metrics-phylogenetic command in QIIME2. At a normalized depth of 4476 reads per sample, three alpha diversity metrics, including the observed features (ASVs), faith's phylogenetic diversity, and shannon diversity index, were calculated for each sample. QIIME2's "alpha rarefaction" command was used to generate rarefaction curves of alpha diversity at different read depths. The "alpha-group-significance" command was utilized to execute the Kruskal-Wallis (KW) test to identify statistically significant differences between the four seasonal groups and the eight geographic groups [2]. Based on the weighted (structure) and unweighted (membership) UniFrac distances, beta diversity was determined. Using the Vegan utility, PCoA graphs based on the weighted and unweighted UniFrac distances were generated. To evaluate the significance of influential factors, the permutation multivariate analyses of variance (PERMANOVA) pseudo-F statistic test based on 999 permutations was implemented [2]. The Benjamin-Hochberg method was used to ascertain significance (adjusted-$p < 0.05$) in alpha and beta diversity assessments when multiple comparisons were performed. Using the arithmetic "stat ellipse" in the R package ggplot2 and the multivariate normal distribution with a 95% confidence interval, seasonal samples were grouped using ellipses. The "dplyr" and "Indicspecies" (v1.7.12) in R was used to identify indicator taxa at the specificity and

sensitivity values of 0.70 for each season and location in respective outdoor communities using triplicate samples [2].

3. Taxonomic Composition of Bacteria in Seasonal Aerosol Samples

3.1. Taxonomic Overview of the Outdoor Air Microbiome at the Phylum, Class, and Genus Levels.

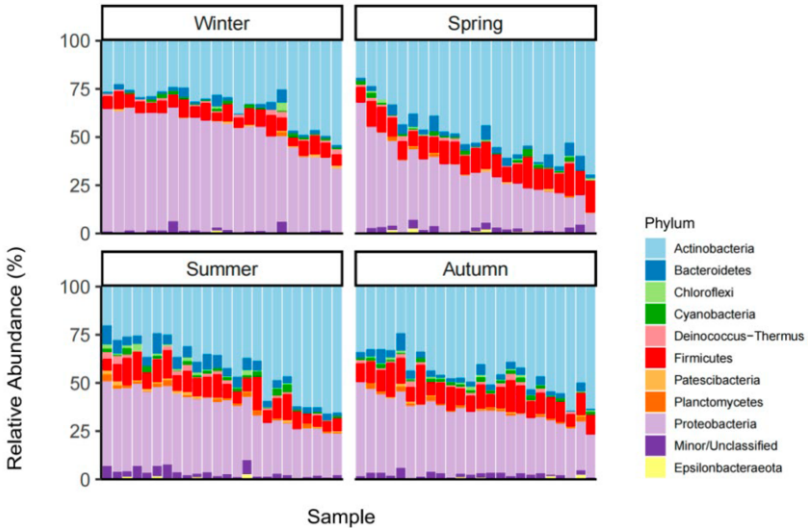

Figure 2. The relative abundance of the outdoor air microbiome in four seasons at the phylum level

In all four seasons, actinobacteria (about 47%) and proteobacteria (about 41%) were the predominant phyla, followed by firmicutes nearly 10% (Figure 2). The relative abundance of proteobacteria was higher than that of actinobacteria in winter, while the opposite pattern was observed in the other three seasons. Other bacterial phylum, such as chloroflexi, cyanobacteria, and epsilonbacteraeota, remained at low levels lower than 5% in all four seasons, suggesting their rarity in outdoor environments. (Figure 1) Overall, the microbial composition of outdoor air was highly similar across four seasons at the phylum level.

At the genus level, cutibacterium dominate all the samples. Cutibacterium is derived from the actinobacteria class, primarily from human skin and sweat, consistent with the dominance of actinomycetota in all the samples (Figure 3). As Hong Kong is one of the most densely populated cities in the world, residents' flaking epidermis may discharge cutibacterium [2]. Halomonas and shewanella dominated the winter samples, reaching nearly 25% and 11% while presented in low abundances or even absent in the other seasons. The remaining microbial populations, such as enhyfrobacter, bacillus and staphylococcus, remained low, with a relative abundance of < 5% in all seasons (Figure 3).

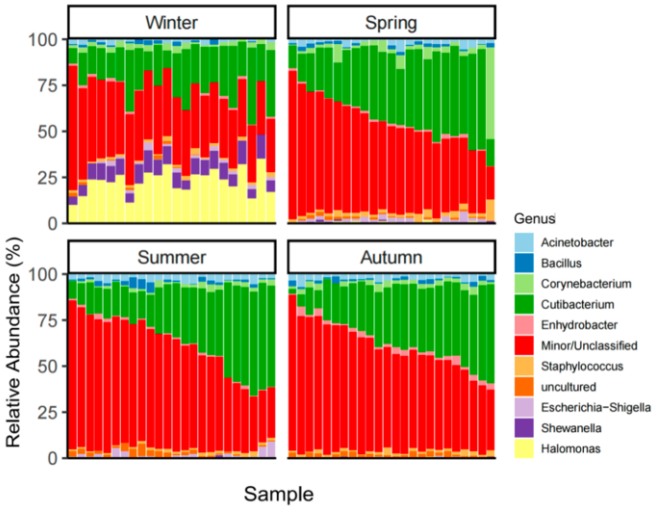

Figure 3. The relative abundance of the outdoor air microbiome in four seasons at the genus level

In general, the quantitative composition of the outdoor air microbial community in Hong Kong was slightly different in winter compared to the high similarity between the spring, summer, and autumn seasons. It showed a higher degree of unevenness. This may be because the natural conditions in winter are more variable in a tropical location like Hong Kong, e.g., temperature and humidity are lower than in other times of the year, and microorganisms such as proteobacteria are more adapted to these natural conditions. In contrast, the growth of some microbial populations, such as actinobacteria, was somewhat suppressed. Still, it is also possible that the decrease in the abundance of bacterial members belonging to actinobacteria, such as cutibacteria, which are adapted to grow on the surface of the skin, is because people go out less in winter due to the low temperature and the reduction in skin exposure.

3.2. Investigating Indicator Species in Outdoor Aerosol Microorganisms during Different Seasons.

Table 1 indicates the species and their corresponding coding (ASV), domain, genus, fidelity, relative abundance, and standard deviation, which occur in autumn, winter, and spring. Indicator analysis identified bacteria that strongly indicate a specific season. EHydrobacter and burkholderiaceae are indicator bacteria for the outdoor air microbiome in autumn, with relative abundance of 0.557%, 0.692%, 0.747%, and 0.452%, respectively.

In autumn, both enhydrobacter and burkholderiaceae occur frequently in aquatic environments, probably associated with Hong Kong being a coastal city. Possible causes of this phenomenon [6], apart from the humidity temperature in autumn, may be since autumn is the change of season in Hong Kong, when the wind direction may change and may blow more air from around the waters towards the urban areas [7]. This may then allow aquatic bacteria to enter the air in the urban area with the air currents. Another reason is that Hong Kong is sometimes affected by typhoons and monsoons in autumn, resulting in increased rainfall. The rain may wash bacteria from the surface and water bodies into the air, thus leading to increased levels of aquatic bacteria in the air.

Table 1. Species and their corresponding codes (ASV), domain, genus, fidelity, relative abundance, standard deviation

Season	ASV	Genus	Fidelity	Average Relative Abundance	Standard Deviation
Autumn	2916	*Enhydrobacter*	1	0.557	0.330
	3f77	*Enhydrobacter*	1	0.692	0.363
	8c68	*Enhydrobacter*	1	0.747	0.356
	c5ec	*Burkholderiaceae*	1	0.452	0.329
Winter	c946	*Shewanella*	1	3.887	1.380
	34f3	*Shewanella*	1	3.886	1.127
	07f8	*Halomonas*	1	4.450	1.419
	8836	*Halomonas*	1	4.387	1.445
	8ab9	*Halomonas*	1	4.386	1.626
	b868	*Halomonas*	1	3.698	1.179
	d38d	*Halomonas*	1	4.526	1.567
Spring	d908	*Corynebacterium*	0.783	1.099	1.567

Shewanella and halomonas strongly indicated winter outdoor microbiome, accounting for 3.887%, 3.886%, 4.450%, 4.387%, 4.386%, 3.698% and 4.526% in the relative abundance. Fidelity levels for the indicator species were 1 in both seasons, with standard deviations of less than 0.400 for the autumn indicator species and less than 1.700 for the winter indicator species, yielding data of high reference value in both seasons (Table 1).

In winter, shewanella is a rod-shaped bacterium that lives mainly in the marine environment [8]. Halomonas belongs to the γ-Amastigotes class of marine halophilic bacteria [9]. A possible reason for them being winter indicator species is that shewanella and halomonas (Table 1), as marine bacteria, may enter the outdoor air with changes in tides, wind direction and marine aerosols. Winter weather conditions and wind direction may contribute to the dispersal of these bacteria in the air. In addition, there may be specific climatic phenomena in winter, such as temperature reversals and air stability, which may affect the distribution and dispersal of airborne bacteria.

Corynebacterium was the only bacterium strongly indicative of spring outdoor microbiome, accounting for 1.099% in the relative abundance. However, the fidelity of the spring indicator species was 0.783 and the standard deviation was also relatively large at 2.390 due to the low availability of reference data (Table 1). Corynebacterium are bacteria that are widely distributed in the natural environment, including soil, water, and air. Some of them are pathogenic to humans, such as corynebacterium diphtheriae and corynebacterium judicium, and may cause diseases of the human immune system [10]. One of the reasons for their appearance in spring may be human activity, as spring temperatures are moderate, outdoor activities increase and people have increased exposure to bacteria in the outdoor air. In addition, spring is the time when plants are flowering and pollen and spore numbers are high, and corynebacterium may attach to these particles and with them enter the outdoor air [11].

3.3. Investigating Influences of Season and Geography on the Outdoor Microbiomes by Alpha-diversity.

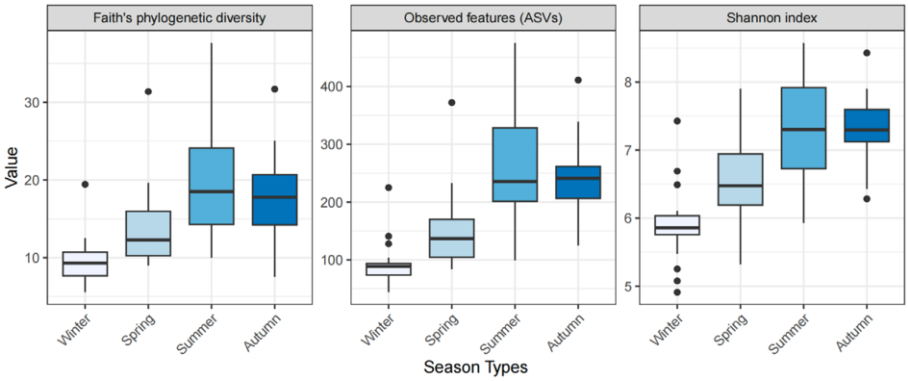

Figure 4. The four seasonal factors that influence the microorganisms in outdoor aerosols were evaluated via three methods: observed features (ASVs), faith's phylogenetic diversity, and shannon diversity index

Before diversity analysis, rarefaction was performed to eliminate the difference caused by the unequal sample reads. Rarefaction analysis revealed that the implemented rarefaction depth could encompass the diversity of the microbial communities (Figure 4). The results showed significant seasonal differences in alpha diversity of outdoor air microbiota were observed between the four seasons regardless of the metrics used (Table 2). The observed difference is the difference observed from two or more data; the critical difference is used to determine whether the observed differences are large enough so that they can be considered statistically significant; and the p-value is used to determine whether the observed differences are statistically significant.

Table 2. Kruskal-Wallis (alpha < 0.05) was used to verify the significance of microbial diversity between seasons

	Observed Difference	Critical Difference	Adjusted-p KW	Statistical Significance
Winter-Spring	21.706	21.346	0.021	TRUE
Winter-Summer	50.686	21.018	0.007	TRUE
Winter-Autumn	49.519	21.018	0.012	TRUE
Spring-Summer	28.980	20.778	0.001	TRUE
Spring-Autumn	27.813	20.778	0.008	TRUE
Summer-Autumn	1.167	20.556	0.312	FALSE

Interestingly, a gradual increase in diversity was observed with time, with the lowest microbial diversity found in winter and the highest diversity in the summer and autumn (adjusted $p = 0.312 > 0.050$, KW), highlighting the critical role of seasonality in governing the dynamics of outdoor air microbial diversity.

4. Investigating Influences of Season and Geography on the Outdoor Microbiomes by Beta-diversity.

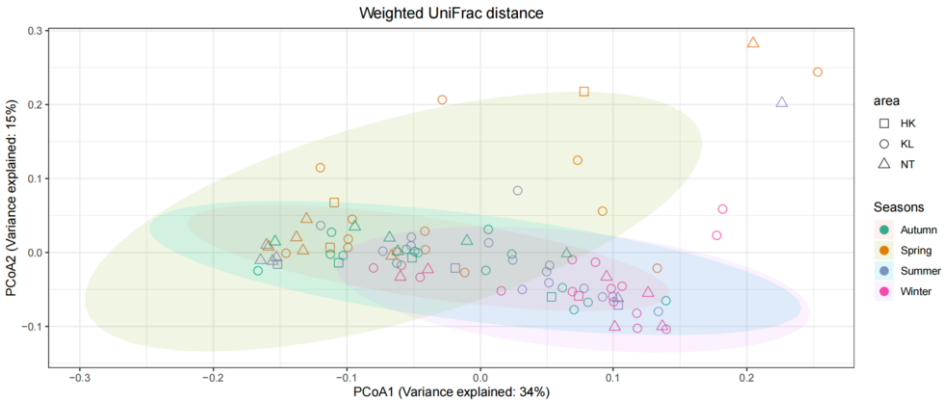

Figure 5. PCoA diagrams of the UniFrac weighted distance of airborne microbiota for the three categories of areas.

The difference in microbial communities between samples was studied using Weighted and Unweighted UniFrac distance metrics and visualized using PCoA. The significance of seasonality and geography on the composition of outdoor airborne microbial communities were investigated using the PERMANOVA test [12]. The heterogeneous dispersion effect had little impact on the PERMANOVA results because the dataset was largely balanced within each group (i.e., each group had nearly the same sample size) when each of the aforementioned factors was considered [2].

As Figure 5 shows, the statistics were calculated using 999 permutations and the PERMANOVA test. The colored ellipses are based on the 95% confidence intervals of the multivariate normal distribution. When the weighted UniFrac distance was considered, the two dimensions explained a total of 49% variance. As can be seen from the graph, the samples in autumn were distinct from that in other seasons for Kruskal-Wallis alpha $p = 0.001 < 0.05$, suggesting that autumn outdoor air microbiomes were different from other seasons in composition. However, the samples in the other three seasons were highly overlapped, suggesting a high similarity in their microbial community composition. Graphical location showed no significant influence on the areas and composition of the outdoor air microbiome (permanova $p = 0.119 > 0.05$, $R^2 = 0.511$, and pseudo-F = 34.428). Taken together, our findings suggested that seasonal change was the primary cause of dynamics in outdoor microbial composition and structure in Hong Kong, with the geographical location only showing a minor effect.

5. Conclusion

Our study highlights the important role of seasonality in shaping the diversity and composition of the Hong Kong outdoor airborne microbiomes. The dominant genera remained stable at each location across the four seasons in outdoor environments, this is likely because Hong Kong's climatic conditions vary within a relatively narrow range despite having four seasons [2]. Nevertheless, indicator bacteria were detected for the three seasons. These findings suggest that environmental conditions are likely to have a

more significant effect on the less abundant members of microbial communities or that some microbial members may be boomed in a specific season. Consistent with this, the beta-diversity analysis demonstrated that the season was a weak but significant driver of the membership and structure of outdoor airborne microbiome.

The community status of outdoor microbial samples fluctuated the most under the influence of seasonal changes, the aerosol microbial diversity peaks in summer and autumn and is lowest in winter, while the influence of regional location was slight than the season factor, suggesting outdoor airborne bacterial communities tended to be homogeneous across different locations on a city-wide scale within a season. By studying the bacterial composition of Hong Kong's outdoor air, we can better understand the role of microorganisms in the urban ecosystem. This will help develop urban planning and environmental protection measures to maintain ecological balance and promote sustainable development.

References

[1] Bolyen, E., Rideout, J. R., Dillon, M. R., & al., e. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature biotechnology, 37(8), 852-857.

[2] Zhou, Y., Lai, Y., Tong, X., & al., e. (2020). Airborne bacteria in outdoor air and air of mechanically ventilated buildings at city scale in Hong Kong across seasons. Environmental Science & Technology, 54(19), 11732-11743.

[3] Prussin, A. J., 2nd, Garcia, E. B., & Marr, L. C. (2015). Total Virus and Bacteria Concentrations in Indoor and Outdoor Air. Environ Sci Technol Lett, 2(4), 84-88. doi: 10.1021/acs.estlett.5b00050

[4] Chen, X., Ran, P., Ho, K., & al., e. (2012). Concentrations and Size Distributions of Airborne Microorganisms in Guangzhou during Summer. Aerosol and Air Quality Research, 12(6), 1336-1344. doi:10.4209/aaqr.2012.03.0066

[5] Zhou, Y., Leung, M. H. Y., Tong, X., & al., e. (2021). City-Scale Meta-Analysis of Indoor Airborne Microbiota Reveals that Taxonomic and Functional Compositions Vary with Building Types. Environmental Science & Technology, 55(22), 15051-15062. doi: 10.1021/acs.est.1c03941

[6] Carrión, V. J., Cordovez, V., Tyc, O., & al., e. (2018). Involvement of Burkholderiaceae and sulfurous volatiles in disease-suppressive soils. The ISME journal, 12(9), 2307-2321.

[7] Staley, J. T., Irgens, R. L., & Brenner, D. J. (1987). Enhydrobacter aerosaccus gen. nov., sp. nov., a gas-vacuolated, facultatively anaerobic, heterotrophic rod. International journal of systematic bacteriology, 37(3), 289-291.

[8] Holt, H., Gahrn-Hansen, B., & Bruun, B. (2005). Shewanella algae and Shewanella putrefaciens: clinical and microbiological characteristics. Clinical microbiology and infection, 11(5), 347-352.

[9] Mata, J. A., Martínez-Cánovas, J., Quesada, E., & al., e. (2002). A detailed phenotypic characterisation of the type strains of Halomonas species. Systematic and applied microbiology, 25(3), 360-375.

[10] Eggeling, L., & Bott, M. (2005). Handbook of Corynebacterium glutamicum: CRC press.

[11] Lee, J.-Y., Na, Y.-A., Kim, E., & al., e. (2016). The actinobacterium Corynebacterium glutamicum, an industrial workhorse.

[12] Jiang, S., Sun, B., Zhu, R., & al., e. (2022). Airborne microbial community structure and potential pathogen identification across the PM size fractions and seasons in the urban atmosphere. Science of The Total Environment, 831, 154665. doi: https://doi.org/10.1016/j.scitotenv.2022.154665